Die objektive Erfassung von Müdigkeit während monotoner Tagfahrten

und deren verbale Selbsteinschätzung durch den Fahrer

Inauguraldissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Eike Andreas Schmidt

geboren in Bielefeld

Juni 2010

Aus dem Institut für Experimentelle Psychologie der
Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Axel Buchner
Korreferent: PD Dr. Michael Schrauf

Tag der mündlichen Prüfung: 09.07.2010

Inhaltsverzeichnis

Zusammenfassung

Zwischen 15 und 20% aller Straßenverkehrsunfälle sind durch Müdigkeit bedingt (Åkerstedt, 2000). Ein bedeutsamer Anteil dieser Unfälle, die oft mit überproportional schweren Folgen einhergehen, ereignen sich am Tage, zumeist unter monotonen Bedingungen und müssen dabei nicht notwendigerweise mit einem Einschlafen des Fahrers verbunden sein (Folkard, 1997; George, 2003; Sagberg, 1999). Zur Vermeidung dieser Unfälle bildet die zuverlässige Selbsteinschätzung des Fahrers eine bedeutsame Fähigkeit, da ein rechtzeitiges Pausieren der Ausbildung von Müdigkeit vorbeugen kann (McDonald, 1984). Der subjektive Zustand von Autofahrern wird zudem häufig als zentrales Müdigkeitsmaß in experimentellen Untersuchungen eingesetzt. Diverse Studien stellen jedoch die Validität der Selbsteinschätzung in Frage oder postulieren eine Beeinflussung des Fahrerzustands durch den Prozess der Abfrage (Kaida et al., 2007; Philip et al., 2003). Durch ihren objektiven Charakter und ihre hohe zeitliche Auflösung eignen sich neurophysiologische Maße, insbesondere die Leistung im EEG-Alpha-Band, hingegen besonders gut für die Erkennung von Fahrermüdigkeit (Lal & Craig, 2001). Ziel der vorgestellten Arbeiten war zum einen die Validierung einer neuen EEG-basierten Methode zur Erfassung von Müdigkeit auf Grundlage von Alpha-Spindeln im Vergleich zu etablierten spektralen EEG-Leistungsmaßen. Des Weiteren war es Ziel, die Zuverlässigkeit subjektiver Maße im Vergleich zu Verhaltens- und physiologischen Maßen zu untersuchen. Dabei war zudem von Interesse, ob die verbale Interaktion im Rahmen der Abfrage den Fahrerzustand beeinflusst und für wie lange ein möglicher Effekt anhält.

Zur experimentellen Bearbeitung der Fragestellungen wurden zwei etwa vierstündige Fahrversuche unter monotonen Tagfahrtbedingungen im Realverkehr durchgeführt. In Studie 1 zeigte sich, dass für Probanden, welche die Fahrt auf Grund starker Müdigkeit abbrachen, die Rate, Amplitude und Dauer von Alpha-Spindeln höhere Effektstärken für die Diskriminierung eines wachen Zustands zu Beginn der Fahrt vom müden Endzustand zeigten als klassische, auf der Leistung in EEG-Bändern basierende Maße. Im Rahmen von Studie 2 wurde belegt, dass nach einer dreistündigen Fahrt eine subjektiv von Probanden berichtete Verbesserung ihres Zustands mit einer weiteren Verschlechterung der objektiven Müdigkeitsindikatoren wie Alpha-Spindelrate, P3-Amplitude, Herzrate und langsame Reaktionszeiten einherging. Studie 3 konnte zeigen, dass eine verbale Interaktion des Versuchsleiters mit dem Fahrer im Rahmen einer Abfrage des subjektiven Zustands in einer kurzfristigen signifikanten Reduzierung der objektiven Müdigkeitsindikatoren Alpha-Spindelrate, Lidschlussdauer, Herzrate und langsamen wie schnellen Reaktionszeiten resultiert. Dem Ende der Abfrage folgend hielt diese Aktivierung für maximal zwei Minuten an.

Zusammenfassend konnte das hohe Potential der EEG-Alpha-Spindelmaße zur Erkennung von Müdigkeit unter Beweis gestellt werden. Die verbale Abfrage des subjektiven Fahrerzustands ist hingegen sowohl unter Aspekten der Validität als auch der Intrusivität des Maßes kritisch zu betrachten. Implikationen der Befunde für die experimentelle Forschung sowie für die praktische Anwendung werden diskutiert.

Abstract

Researchers claim that between 15 and 20% of all traffic accidents are caused by fatigue and that these accidents often result in disproportionately severe consequences (Åkerstedt, 2000; Sagberg, 1999). A significant portion of fatigue related accidents occur at daytime, mostly under monotonous conditions and do not have to be accompanied by the driver falling asleep (Folkard, 1997; George, 2003). In order to prevent these accidents the drivers' ability to assess their own state is of crucial importance since resting can prevent the course of fatigue (McDonald, 1984). In addition the subjective state is an often-applied measure in experimental studies on driver fatigue. A fair amount of studies question the validity of self assessment measures or suspect that the process of asking the driver is influencing the state being under investigation (Kaida et al., 2007; Philip et al., 2003). In contrast, due to their objectivity and their high temporal resolution, neurophysiological measures and especially EEG alpha power are well suited for the assessment of driver fatigue (Lal & Craig, 2001). The studies presented here had the objective of validating a new EEG-based method for the assessment of driver fatigue. This method is based on alpha spindles and is compared to established EEG power measures. Further the reliability of subjective measures as compared to performance measures and physiological measures was to be investigated. It was of particular interest if the verbal interaction during fatigue assessment influenced driver state and for how long this effect might last.

In order to address these issues two four hour driving studies were conducted under monotonous daytime conditions in real traffic. In study 1 for drivers who aborted the drive due to heavy fatigue higher effect sizes for discriminating states of wakefulness and fatigue were observed for the dependent measures alpha spindle rate, alpha spindle amplitude and alpha spindle duration as compared to classic EEG measures solely based on spectral power. In study 2 it could be shown that after three hours of driving a subjectively reported improvement of vigilance state was accompanied by a further degradation of objective fatigue measures, e.g. alpha spindle rate, P3 amplitude, heart rate and slow reaction times. Study 3 proved that a verbal interaction between driver and investigator in the course of a subjective state assessment resulted in a significant reduction of objective fatigue measures, e.g. alpha spindle rate, eye blink duration, heart rate as well as slow and fast reaction times. This activation persisted for up to two minutes following the end of the interaction.

In conclusion the high potential of EEG alpha spindle measures for the detection of driver fatigue was shown. In contrast the verbal assessment of driver state has to be applied with caution due to issues of validity and intrusion. Implications of these findings for the experimental study of driver fatigue and their real life applications are discussed.

1 Einleitung

In den EU-Staaten starben im Jahr 2007 etwa 39.000 Menschen im Straßenverkehr (CARE, 2009). Dabei stellt der Mensch selbst als Verursacher von ca. 90 % aller Unfälle den mit Abstand häufigsten Auslöser dar (DESTATIS, 2007). Obwohl Müdigkeit als Unfallursache schwer nachzuweisen ist und deshalb in den offiziellen Statistiken nicht adäquat erfasst werden kann, kommt eine Vielzahl von Studien zu dem Schluss, dass 15 bis 20% aller Verkehrsunfälle durch Müdigkeit bedingt sind (Åkerstedt, 2000; MacLean, David & Thiele, 2003). Wissenschaftler gehen davon aus, dass Müdigkeit damit eine häufigere Unfallursache darstellt als der Konsum von Alkohol oder Drogen (Åkerstedt, 2000).

Im Gegensatz zur Mehrzahl der Studien zum Thema Müdigkeit im Straßenverkehr, die zumeist im Simulator, mit schlafdeprivierten Probanden und oftmals unter nächtlichen Bedingungen durchgeführt wurden, war es Ziel der vorliegenden Arbeiten, die Auswirkungen monotoner *Tagfahrten* im *realen Straßenverkehr* und mit *ausgeschlafenen Fahrern* zu untersuchen. So ist bekannt, dass sich Entwicklung und Ausmaß von Müdigkeit unter Realfahrtbedingungen und im Simulator deutlich unterscheiden (Belz, Robinson & Casali, 2004; Philip et al., 2005). Erkenntnisse aus Fahrversuchen im realen Straßenverkehr stellen somit eine höhere ökologische Validität sicher. Ein methodischer Schwerpunkt wurde auf die Evaluation und den Einsatz des Elektroenzephalogramms zur Erfassung von Fahrermüdigkeit gelegt, da diese Methode durch die direkte Abbildung kognitiver Prozesse mit einer hohen zeitliche Auflösung (Lal & Craig, 2001; Lal & Craig, 2002; Lin, Wu, Jung, Liang & Huang, 2005) ein hohes Potential zur präzisen Erfassung von Ermüdungsvorgängen trägt. Von zentraler Bedeutung waren zudem die Zuverlässigkeit und die Anwendbarkeit der Selbsteinschätzung der Müdigkeit durch den Fahrer. Diese Thematik gewinnt seine besondere Relevanz vor allem dadurch, dass eine zuverlässige Selbsteinschätzung die notwendige Bedingung für ein rechtzeitiges Pausieren oder einen Fahrerwechsel bildet (McDonald, 1984).

1.1 Theoretischer Hintergrund

Obwohl der Begriff Müdigkeit in seiner alltäglichen Verwendung eindeutig erscheint, wird er bei einem genaueren Blick in die einschlägige Fachliteratur durchaus uneinheitlich verwendet (Radun, 2009). Deshalb soll im Folgenden die am Modell von May und Baldwin (2009) orientierte und in den vorliegenden Arbeiten verwendete Terminologie näher erläutert werden. May und Baldwin unterscheiden auf Grundlage der bedingenden Faktoren zwischen schlafbezogener (sleep-related) und aufgabenbezogener (task-related) Müdigkeit (fatigue), wobei aufgabenbezogene Müdigkeit nochmals in die Kategorien „aktiv“ und „passiv“ unterteilt wird, je nachdem ob die ausgeführte Tätigkeit in hohem oder geringem Maße beanspruchend ist (Desmond & Hancock, 2001; Tejero Gimeno, Pastor Cerezuela & Chóliz Montañés, 2006).

Schlafbezogene Müdigkeit, von vielen Autoren auch als Schläfrigkeit (sleepiness) bezeichnet, wird vorwiegend durch Schlafmangel und den zirkadianen Rhythmus beeinflusst. Die Bedeutung schlafbezogener Faktoren für den Straßenverkehr verdeutlichen zum Beispiel relativ erhöhte Unfallraten in den frühen Morgenstunden und am Nachmittag, also während zirkadianer Phasen erhöhter schlafbezogener Müdigkeit (Folkard, 1997) oder auch erhöhte Unfallzahlen für Autofahrer mit Schlafstörungen (z.B. Obstruktives Schlafapnoe Syndrom: George, Nickerson, Hanly, Millar & Kryger, 1987). Schlafbezogene Müdigkeit kann definiert werden als die Schwierigkeit wach zu bleiben und kann nur durch Schlaf, nicht jedoch durch Ruhephasen reduziert werden (Philip et al., 2005). Die physiologischen Hintergründe des menschlichen Schlafes diskutieren Dijk und von Schanz (2005) in einer Übersichtsarbeit.

Aufgabenbezogene Müdigkeit wird hingegen durch externe Einflussfaktoren bedingt, tritt in der Regel erst nach einer gewissen Aufgabendauer auf und steigt mit zunehmender Bearbeitungszeit an. *Passive aufgabenbezogene Müdigkeit* wird dabei durch eine stark monotone Aufgabe mit sehr geringer Beanspruchung und hoher Vorhersehbarkeit provoziert (Wertheim, 1991), wie es auf wenig befahrenen Straßen in reiz- und ereignisarmer Umgebung der Fall ist. Einige Autoren befürchten, dass die zunehmende Automatisierung durch Fahrerassistenzsysteme zu einer Erhöhung der passiven aufgabenbezogenen Müdigkeit führen könnte (Hancock & Verwey, 1997). Die *aktive aufgabenbezogene Müdigkeit* wird hingegen durch eine hohe Beanspruchung ausgelöst, wie sie bei erhöhtem Verkehrsaufkommen, anspruchsvollen Manövern, schlechter Sicht oder durch Nebentätigkeiten während der Fahrt ausgelöst wird (Tejero Gimeno et al., 2006). Passiver und aktiver aufgabenbezogener Müdigkeit kann durch Pausieren, also das zeitweilige Unterbrechen der Aufgabe entgegengewirkt werden (Philip et al., 2005).

Darüber hinaus finden sich im Zusammenhang mit monotonem Autofahren häufig die Begriffe „drowsiness" („Benommenheit") und „vigilance [decrement]" (Vigilanz[abfall]). Diese Begrifflichkeiten werden im Rahmen des erweiterten Modells als Folgen von Müdigkeit eingeordnet (Abbildung 1). „Drowsiness" bezeichnet dabei allgemeingebräuchlich eher den subjektiven oder physiologischen Zustand reduzierter Aufmerksamkeit (Johns, 2000), während Vigilanz vorwiegend über Leistungseinbußen in der Fähigkeit auf Reize in der Umwelt zu reagieren definiert wird (Mackworth, 1957; Parasuraman, 1998). Der Begriff „Highway Hypnosis" beschreibt einen Zustand, der ebenfalls in Verbindung mit monotonem Autofahren diskutiert wird (Wertheim, 1991). Dieser Zustand tritt laut Definition in einer hoch vorhersehbaren Umgebung auf, kann von Müdigkeit begünstigt werden und auch bei ausgeruhten Fahrern auftreten. Als zentrale Auswirkungen werden eine Reduzierung des Situationsbewusstseins sowie eine Beeinträchtigung der Fahrleistung genannt.

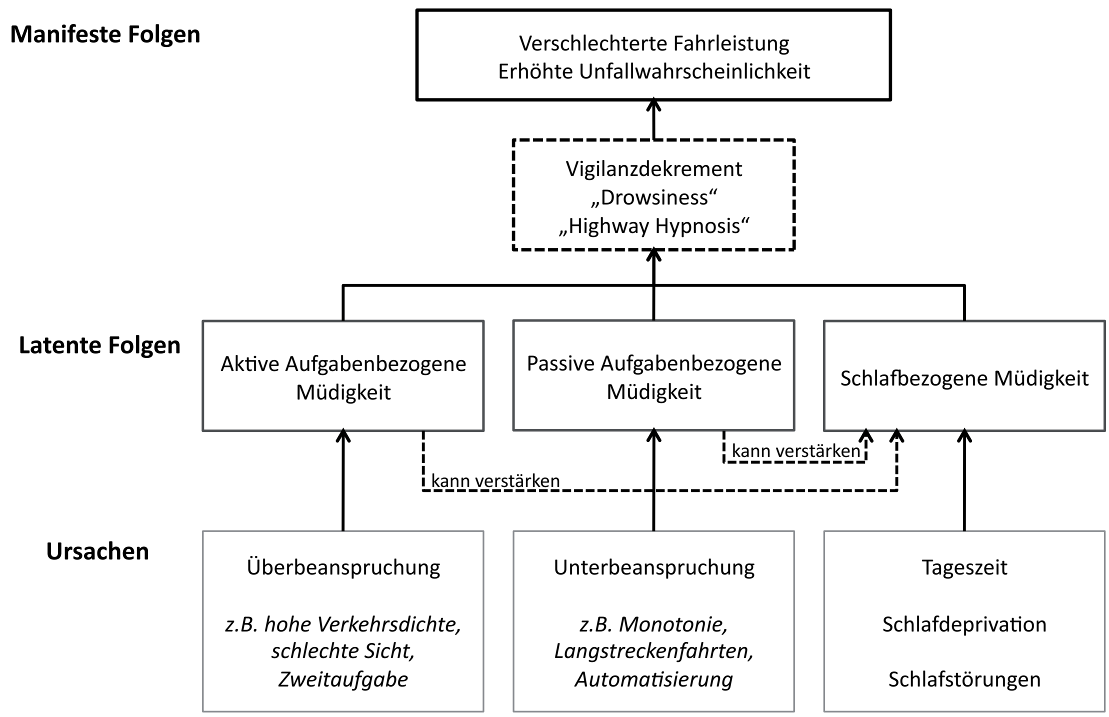

Abbildung 1: Erweitertes Müdigkeitsmodell nach May und Baldwin (2009).

Abbildung 1 gibt einen Überblick über die ursächlichen Zusammenhänge der oben erläuterten Begrifflichkeiten. Tabelle 1 zeigt die Ergebnisse einer Abfrage der Datenbank PsycINFO (American Psychological Assosciation) über die gemeinsame Nennung der oben erläuterten Konstrukte in Verbindung mit dem Thema Autofahren („driving“). Die Abfrage belegt, dass die Begriffe „fatigue“, „sleepiness“ und „vigilance“ in ähnlichen Häufigkeiten verwendet werden und es somit bislang keine klare Tendenz für die vermehrte Verwendung eines bestimmten Begriffs gibt.

Tabelle 1: PsycINFO (APA) Abfrage vom 10.01.2010: Anzahl der Treffer der einzelnen Begriffe in Verbindung mit „driving“.

	Alle Felder	Abstract	Schlüsselwörter
Fatigue	1068	193	82
Sleepiness	824	161	58
Vigilance	836	103	29
Drowsiness	216	49	16
Highway hypnosis	21	2	2

Ziel der im Folgenden näher beschriebenen eigenen Arbeiten war es, den Einfluss monotoner Tagfahrten und somit passiver aufgabenbezogener Müdigkeit möglichst realitätsnah zu untersuchen. Allerdings ist es nicht möglich, Einflüsse schlafbezogener Müdigkeit gänzlich auszuschließen. Zwar schliefen die Probanden nach eigenen Angaben stets normal in der Nacht vor der Versuchsfahrt und es wurde durch den Versuchsplan keine Manipulation der schlafbezogenen Müdigkeit vorgenommen, jedoch wurden die Fahrten im zirkadianen Nachmittagstief durchgeführt und die Schlafhistorie nur für die vorangegangene Nacht erfasst. Zirkadiane Einflüsse oder ein mögliches kumulatives Schlafdefizit können deshalb nicht gänzlich ausgeschlossen werden.

May und Baldwin (2009) führen an, dass sich die Auswirkungen und damit die auch die Messmethoden für die unterschiedlichen Müdigkeitsformen gleichen. Diese Methoden werden in Abschnitt 1.3 vorgestellt. Im Gegensatz dazu sollten sich die Gegenmaßnahmen in Abhängigkeit der Müdigkeitsform voneinander unterscheiden. Dieser Aspekt wird in Abschnitt 1.4 aufgegriffen.

1.2 Müdigkeit und Autofahren

Autofahren lässt sich als komplexe, aber dennoch hoch automatisierte Aufgabe beschreiben. Michon (1985) unterteilt den Prozess des Autofahrens in drei Aktivitätsebenen: die *strategische Ebene*, die *taktische Ebene* und die *Kontrollebene.* Auf der strategischen Ebene werden Entscheidungen über die allgemeinen Ziele der Fahrt getroffen (z.B. Startzeit, Fahrtroute). Auf der taktischen Ebene fallen Entscheidungen über die unmittelbaren Ziele (z.B. Geschwindigkeit, Fahrspur, Manöver). Die Kontrollebene schließlich ist bestimmt durch die augenblickliche Steuerung des Fahrzeugs, bei der es essentiell ist, Veränderungen in der Umwelt fortlaufend zu detektieren und auf diese adäquat zu reagieren (z.B. Längs- und Querführung). Alle drei Ebenen können durch Müdigkeit negativ beeinflusst werden. Eine direkte Gefahr stellen dabei Leistungseinbußen auf der Kontrollebene dar, weil diese durch ein Verlassen der Spur oder das Auffahren auf den Vordermann unmittelbar in einem Unfall resultieren können. Müdigkeit erzeugt laut Brown (1994) eine sukzessive Rücknahme der Aufmerksamkeit von der Straße. Somit ist auch ohne ein Einschlafen des Fahrers mit einer erhöhten Unfallgefahr in Folge von Müdigkeit zu rechnen (George, 2003).

Die Mehrheit der Unfallforscher vertritt die Meinung, dass offizielle Statistiken Müdigkeit als Unfallursache stark unterschätzen (Åkerstedt, 2000). Diese Annahme wird in erster Linie dadurch begründet, dass die Erfahrung von Müdigkeit vor einem Unfall in anonymen Befragungen deutlich häufiger benannt wird, als dies aus offiziellen Statistiken zu erkennen ist. Verschiedene Erklärungen werden als Ursache für diese Diskrepanz herangezogen: Einerseits wird in den meisten Ländern Müdigkeit in den offiziellen Formblättern zur Erfassung der Unfallursache nicht aufgeführt (Horne & Reyner, 1999), andererseits kann die Erinnerung für die Erfahrung von Müdigkeit vor dem Unfall durch traumainduzierte oder reguläre Vergessensprozesse getrübt sein beziehungsweise auch die Erinnerung bewusst verschwiegen werden, da negative Konsequenzen versicherungsoder strafrechtlicher Art befürchtet werden. Darüber hinaus ist es Polizeibeamten nicht zweifelsfrei möglich trotz eines abweichenden Berichts des Fahrers, Müdigkeit als Unfallursache zu erkennen und diese bei einer juristischen Auseinandersetzung zu beweisen - auch dadurch bedingt, dass es keine klaren Kriterien für diese Klassifizierung gibt (Radun, 2009).

Unfälle, die ursächlich im Zusammenhang mit Müdigkeit stehen, sind überproportional häufig mit Tod oder schweren Verletzungen verbunden (Horne & Reyner, 1999). Dies ist vorwiegend durch hohe Geschwindigkeiten, seitliches Verlassen der Fahrbahn und das Ausbleiben beziehungsweise deutlich zu späte Einleiten von Ausweich- beziehungsweise Bremsmanövern bedingt (Pack et al., 1995; Radun, 2009). Sagberg (1999) identifizierte eine monotone Umgebung, etwa in Form einer geraden Fahrbahngeometrie und eine niedrige Verkehrsdichte als bedeutsame externe Ursachen für müdigkeitsbedingte Unfälle. Einen validen Beleg fand diese These durch den Befund, dass sich in Neuseeland 80% aller müdigkeitsbedingten Unfälle in Bereichen mit sehr niedrigen Anforderungen ereignen (Smith, Oppenhuis & Koorey, 2006).

Vergleicht man die häufig diskutierten und in verbindlichen Grenzwerten umgesetzten Auswirkungen von Alkohol mit denen von Müdigkeit, so zeigt sich, dass etwa 18 Stunden Wachheit in visuomotorischen Ausfällen gleichen Ausmaßes resultieren wie eine Blutalkoholkonzentration von 0,5 Promille (Arnedt, Wilde, Munt & MacLean, 2001; Dawson & Reid, 1997).

1.3 Methoden zur Erfassung von Müdigkeit beim Autofahren

Beim Begriff Müdigkeit handelt es sich um ein hypothetisches, psychologisches Konstrukt, das einer Operationalisierung bedarf (Popp, 2005). Klassisch wird zwischen einer Operationalisierung mittels subjektiver Maße (Selbsteinschätzung), mittels verhaltensbezogener Maße (Leistung in einer Aufgabe oder Verhaltensbeobachtung) und mittels physiologischer Maße unterschieden. Dabei wird entweder das spontane Verhalten beziehungsweise der spontane physiologische Zustand oder aber ein durch einen Reiz evoziertes Verhalten beziehungsweise eine entsprechende physiologische Reaktion erfasst.

Der Kontext des Autofahrens stellt besondere Anforderungen an Methoden zur Erfassung von Müdigkeit. Neben den notwendigen Bedingungen Sensitivität, Spezifität und Reliabilität bilden in diesem Kontext insbesondere die Einsetzbarkeit während des Fahrens (Mobiliät) sowie eine gute Benutzbarkeit (usability) zentrale Anforderungen. Zudem sollte die Messung nicht den zu untersuchenden Zustand beeinflussen, das heißt, weder Müdigkeit induzieren noch zu ihrer Reduzierung beitragen (geringe Intrusivität). Die zudem wichtige Anforderung einer hohen zeitlichen Auflösung stellt schon deshalb eine große Herausforderung dar, weil der tatsächliche Müdigkeitszustand niemals eindeutig bekannt ist: Es fehlt die so genannte „ground truth“. Zwar kann zum Beispiel durch Schlafdeprivation oder wie in Studie 1 durch lang andauerndes Autofahren ein Zustand erzeugt werden, welcher bestimmten Kriterien zur Operationalisierung von Müdigkeit entspricht, jedoch finden sich auch innerhalb dieses Zustands deutliche Schwankungen. So ist Müdigkeit kein streng monotoner Prozess, sondern kann mit hoher Frequenz fluktuieren und dies insbesondere dann, wenn gegen die Müdigkeit angekämpft wird (Lal & Craig, 2005).

Die zuverlässige Induktion und Erfassung von Müdigkeit wird zusätzlich dadurch erschwert, dass bedeutsame Unterschiede zwischen verschiedenen Fahrern in der Anfälligkeit für Müdigkeit und der Fähigkeit mit dieser umzugehen existieren (Desmond & Matthews, 2009). So konnten Verwey und Zaidel (2000) sowie Thiffault und Bergeron (2003b) zeigen, dass die Persönlichkeitseigenschaften Extraversion beziehungsweise Sensation Seeking einen deutlichen Einfluss auf die Fahrleistung von Probanden unter Müdigkeitsbedingungen hatten.

Auf Grund einer bedeutsamen Varianz in der Sensitivität verschiedener Messmethoden sowohl innerhalb als auch zwischen Individuen (Curcio, Casagrande & Bertini, 2001; Ingre, Åkerstedt, Peters, Anund & Kecklund, 2006a) wurde in den vorliegenden Erhebungen ein Mehrebenenansatz gewählt. Es spricht für eine hohe Validität der Befunde, wenn verschiedene Messmethoden für die Erfassung der gleichen latenten Variable zu konvergenten Ergebnissen gelangen (Moller, Kayumov, Bulmash, Nhan & Shapiro, 2006).

Die Vor- und Nachteile der verschiedenen Messmethoden zur Erfassung von Müdigkeit sowie Anwendungsbeispiele im Kontext des Autofahrens werden in den folgenden Abschnitten erörtert.

1.3.1 Subjektive Maße

Wie eingangs erwähnt, ist unter dem Gesichtspunkt der Verkehrssicherheit die Fähigkeit zur Einschätzung der eigenen Müdigkeit beim Führen eines Fahrzeugs von großer Bedeutung (McDonald, 1984). Im wissenschaftlichen Kontext ist die Befragung des Probanden nach seinem Zustand die „intuitivste" und einfachste Methode zur Erfassung von Müdigkeit. Zu diesem Zweck wurden eine Reihe von Müdigkeitsskalen entwickelt, die sich danach unterscheiden lassen, ob sie die situative Müdigkeit zu einem bestimmten Zeitpunkt erfassen (z.B. Karolinska Sleepiness Scale: KSS, Åkerstedt & Gillberg, 1990; Kaida et al., 2006; Stanford Sleepiness Scale: SSS, Hoddes, Zarcone, Smythe, Phillips & Dement 1973; Visual Analogue Scales: VAS, für eine Übersichtsarbeit siehe Monk, 1987) oder eine generelle Einschlafneigung im Sinne einer Disposition (z.B. Epworth Sleepiness Scale: ESS, Johns, 1991). Diese Skalen wurden unter Verwendung von physiologischen sowie Leistungsmaßen validiert und wiederholt im Kontext des Autofahrens eingesetzt (z.B. SSS: Arnedt et al., 2005; Ting, Hwang, Doong & Jeng, 2008; KSS: Biggs et al., 2007; Ingre et al, 2006a, 2006b; ESS: Jackson et al., 2008). Da die vorliegenden Arbeiten sich mit der situativen Erfassung von Müdigkeit beschäftigen, wurde in allen Studien die KSS eingesetzt.

In der Fachliteratur existieren uneinheitliche Befunde darüber, wie gut ausgeprägt die Fähigkeit eines Fahrers zur Einschätzung seines Zustands beziehungsweise seiner Leistungsfähigkeit unter Müdigkeitsbedingungen ist. Eine Reihe von Autoren berichten, dass die Selbsteinschätzung keinen hinreichend validen Prädiktor für Reaktionszeiten, Fahrleistung oder die Wahrscheinlichkeit einzuschlafen bildet (Belz et al., 2004; Lenné, Triggs, & Redman, 1997; Moller et al., 2006; Philip et al., 1997; Philip et al., 2005). So fanden Philip et al. (2003), dass in Folge von Langstreckenfahrten am Tage sowohl bei schlafdeprivierten als auch ausgeschlafenen Fahrern die Selbsteinschätzung der Reaktionszeit mangelhaft war. Andere Wissenschaftler führen hingegen an, dass die Mehrheit der Fahrer in der Tat eine sehr gute Fähigkeit besitzen, den eigenen Müdigkeitszustand zu erkennen und daraus resultierende Leistungseinbußen einzuschätzen (Baranski, 2007; Horne & Baulk, 2004; Lisper, Laurell, & van Loon, 1986). So berichten Nordbakke und Sagberg (2007), dass Fahrer zwar die eigene Müdigkeit korrekt wahrnehmen und auch ein ausgeprägtes Bewusstsein für die Gefahren von Müdigkeit besitzen, sie ihre Fahrt aber dennoch fortsetzen. Auffällig ist, dass Studien, die eher optimistisch über die Fähigkeiten zur Selbsteinschätzung urteilen, vorwiegend mit schlafdeprivierten Probanden und bei Nacht durchgeführt wurden, also unter Bedingungen, in denen schlafbezogene Müdigkeit den dominierenden Faktor darstellt. Von daher kann vermutet werden, dass Fahrer womöglich ein höheres Bewusstsein für schlafbezogene als für aufgabenbezogene Müdigkeit haben.

Auch wenn dies in der Praxis einen sehr ungewöhnlichen Fall darstellen sollte, kann im Falle subjektiver Maße auch eine willentliche Beeinflussung des Urteils durch den Probanden nicht ausgeschlossen werden. Neben dieser mangelnden Objektivität stellt die Beeinflussung des zu erfassenden Zustands durch den Akt der Befragung einen weiteren Problembereich in Bezug auf die subjektiven Maße dar. Im Falle der KSS erfolgt die Abfrage zumeist schriftlich anhand eines Fragebogens oder aber verbal. Die Häufigkeit der Abfrage ist nicht festgelegt, liegt aber zumeist zwischen einem Intervall alle 3 Minuten bis hin zu Einzelerfassungen (Kecklund et al., 2006). Um die Interferenz mit dem Zustand und den möglicherweise gleichzeitig ausgeführten Aufgaben so gering wie möglich zu halten, ist es anzustreben, eine möglichst niedrige Abfragerate zu wählen. Auf der anderen Seite verringert sich damit die zeitliche Auflösung des Maßes und mögliche Gedächtniseffekte können im retrospektiven Urteil an Einfluss gewinnen. Für die KSS konnten Kaida, Åkerstedt, Kecklund, Nilsson und Axelsson (2007) zeigen, dass deren fünfmalige Abfrage während einer vierzigminütigen Vigilanzaufgabe die subjektive sowie die mittels der Leistung im EEG-Alpha-Band erfasste physiologische Müdigkeit nach Ende der Aufgabe reduzierte. Des Weiteren ist bekannt, dass bestimmte Erfassungskontexte, zum Beispiel eine soziale Interaktion oder körperliche Aktivität, den Zusammenhang zwischen subjektivem Urteil und objektiven Maßen reduzieren können (Åkerstedt, Kecklund & Axelsson, 2008; Eriksen, Åkerstedt, Kecklund & Åkerstedt, 2005; Matsumoto, Mishima, Satoh, Shimiziu & Hishikawa, 2002).

Die beschriebenen Kontroversen in Bezug auf den Einsatz subjektiver Maße waren Anlass, diese genauer zu untersuchen. Studie 2 adressiert dabei die Frage der Selbsteinschätzungsfähigkeit des Fahrers, Studie 3 hingegen den Einfluss der verbalen Abfrage auf den Fahrerzustand.

1.3.2 Leistungsmaße

Grobe Fahrfehler können die Ursache für das Verunfallen eines Autofahrers sein und weisen somit eine hohe Augenscheinvalidität für die Erfassung von Müdigkeit auf. Allerdings kann ein solches Maß offensichtlich nicht als kontinuierlicher Müdigkeitsindikator dienen, da schwere Fahrfehler und Unfälle, insbesondere außerhalb eines Fahrsimulators, sehr seltene Ereignisse darstellen und erst in späten Müdigkeitsstadien auftreten. Deshalb bedarf es subtilerer Leistungsmaße, die sich bereits in frühen Stadien der Müdigkeit sensitiv zeigen. Dabei stellt die Leistung in der Erstaufgabe, also der Fahrzeugführung, zunächst den plausibelsten Ansatzpunkt dar (für eine Übersicht siehe: Liu, Hosking & Lenné, 2009). Häufig in der Literatur berichtete Maße sind dabei die Standardabweichung der lateralen Position in der Spur (z.B. Ingre et al., 2006a), die Zeit bis zur Überquerung des Fahrbahnrandes (z.B. Ting et al., 2008), Lenkbewegungen (z.B. Thiffault & Bergeron, 2003a), Geschwindigkeit (z.B. Lenné et al., 1997) oder der Abstand zum vorausfahrenden Fahrzeug (z.B. Ting et al., 2008). Die Sensitivität von Fahrleistungsmaßen für die Erkennung früher Auswirkungen von Müdigkeit wird allerdings des Öfteren in Frage gestellt. So kann ein Fahrer womöglich noch sehr gut die Spur halten, obwohl seine Reaktionsfähigkeit auf ein unerwartetes Ereignis bereits herabgesetzt ist. Ein zweiter Kritikpunkt ist, dass mögliche Einbußen in der Fahrleistung nicht immer eindeutig Effekten von Müdigkeit zugeordnet werden können. So können die gleichen Ausfallmuster durchaus auch durch Ablenkung oder Phasen hoher Beanspruchung provoziert werden. Zur Auflösung dieser Uneindeutigkeit kann der zusätzliche Einsatz physiologischer Maße beitragen, da diese zwischen Zuständen starker Beanspruchung und Müdigkeit unterscheiden können. Zu guter Letzt gibt es in der Fahrzeugführung bedeutsame Unterschiede zwischen verschiedenen Fahrern, so dass Liu et al. (2009) schlussfolgerten, dass zur zuverlässigen Erkennung von Müdigkeit für eine Vielzahl von Individuen auch die Analyse einer Vielzahl von Fahrerleistungsmaßen nötig ist.

Um eine Sensitivität des verwendeten Leistungsmaßes auch in früheren Müdigkeitsstadien sicherzustellen, wurde in den vorliegenden Studien im Einklang mit O'Donnel und Eggemeier (1986) eine Zweitaufgabe eingeführt. Dem liegt die Annahme zu Grunde, dass die Leistung in einer Zweitaufgabe von den verbliebenen freien Ressourcen bei Durchführung der Erstaufgabe abhängt. Da Müdigkeit in einer Reduzierung der insgesamt verfügbaren Ressourcen resultiert, sollte sich dies zunächst in einer verschlechterten Leistung in der Zweitaufgabe auswirken (Lenné et al., 1997). Den am ausführlichsten validierten und am häufigsten eingesetzten Test zur Erfassung der Vigilanzleistung stellt der Psychomotor Vigilance Task dar (PVT: Dinges & Powell, 1985). Für unseren Anwendungsfall schien der PVT allerdings ungeeignet, da der Müdigkeitszustand *während* der Fahrt erfasst werden sollte und vom PVT als visueller Aufgabe ein zu hoher Grad an Interferenz mit der ebenfalls vorwiegend visuellen Fahraufgabe zu erwarten war. Aus diesem Grund wurde eine auditorische Zweitaufgabe eingesetzt. Laurell und Lisper (1978) konnten im Rahmen einer Realfahrtstudie auf einer abgesperrten Teststrecke einen Zusammenhang zwischen Reaktionszeiten auf einen auditorischen Reiz und Bremsreaktionszeiten auf ein reales Hindernis zeigen. Allerdings wird die Validität von Reaktionszeiten zur Erfassung von Müdigkeit vereinzelt in Zweifel gezogen. So fanden Baulk, Reyner und Horne (2001) in einer monotonen Simulatorstudie, dass Reaktionszeiten in einer Zweitaufgabe keinen validen Indikator für Müdigkeit darstellten, sondern diese Aufgabe lediglich die Monotonie reduzierte.

1.3.2.1 Lange Reaktionszeiten

In einer klassischen Schlafentzugsstudie im militärischen Kontext berichten Williams, Lubin und Goodnow (1959), dass sich der Mittelwert der 10 längsten Reaktionszeiten in einer visuellen Aufgabe sensitiv gegenüber müdigkeitsinduzierten Vigilanzeinbußen zeigte, während das Mittel der kürzesten 10 Reaktionszeiten über den Verlauf unverändert blieb. Aus diesen Befunden leiteten die Autoren die sogenannte „Lapse Hypothesis“ ab, welche mit abnehmendem Vigilanzzustand eine Zunahme von Dauer und Häufigkeit solcher Phasen annimmt, in denen nicht oder nur stark verzögert reagiert wird. Obwohl es verschiedene Befunde gab, welche die „Lapse Hypothesis“ in Zweifel zogen (Lim & Dinges, 2008; Murray, 1965), haben sich dennoch die langsamen Reaktionszeiten in einer Reihe von Studien als guter Indikator für Müdigkeit erwiesen (Graw, Kräuchi, Knoblauch, Wirz-Justice & Cajochen, 2004).

1.3.3 Physiologische Maße

Physiologische Maße zeigen die größten Vorteile in ihrer Objektivität (nicht bewusst beeinflussbar), der geringen Intrusivität (sehr wenig bis keine Interferenz mit der Fahraufgabe) sowie einer hohen zeitlichen Auflösung.

Klassische Maße zur Erfassung der Müdigkeit auf Grundlage des EEGs sind der Multiple Sleep Latency Test (MSLT: Carskadon et al., 1986) und der Maintenance of Wakefulness Test (MWT: Mitler, Gujavarty & Browman, 1982). Sie erfassen Müdigkeit anhand der Zeit, die Probanden bis zum mittels des EEGs definierten Einschlafen benötigen (MSLT) beziehungsweise die sie wach bleiben können (MWT). Diese Methoden stellen zwar eine valide Operationalisierung von Müdigkeit dar, jedoch kommen sie im vorliegenden angewandten Kontext nicht in Frage, da eine Unterbrechung der Fahraufgabe zu ihrer Erfassung nötig wäre.

Eine Auswahl an physiologischen Maßen, die sich in verschiedenen Kontexten sensitiv für Müdigkeit gezeigt haben, sind das EEG (Spontan-EEG: Lal & Craig, 2001; Oken, Salinsky & Elsas, 2006 [siehe Abschnitt 1.3.3.1 für Alpha-Band] und ereigniskorrelierte Potentiale [siehe Abschnitt

1.3.3.2]), mittels Elektrookulogramm (EOG) oder Bildverarbeitung erfasste Lidschlussmaße (siehe Abschnitt 1.3.3.4.) und Augenbewegungen (z.B. Schleicher, Galley, Briest & Galley, 2008), Pupillographie (Wilhelm, 2008; Yoss, 1969), Elektrokardiographie (EKG: Herzrate siehe Abschnitt 1.3.3.3; Variabilitätsmaße siehe z.B. Tran, Wijesuriya, Tarvainen, Karjalainen & Craig, 2009), Hautleitfähigkeit (z.B. Wright & McGown, 2001), Elektromyographie (EMG: z.B. Dureman & Bodén, 1972), Atemfrequenz (z.B. Fairclough & Venables, 2006), Melatonin-Level (z.B. Phipps-Nelson, Redman, Schlangen & Rajaratnam, 2009) und die Flimmer-Verschmelzungs-Frequenz (z.B. Iudice et al., 2005). Für eine Übersicht und Evaluation verschiedener physiologischer Maße zur Erkennung von Müdigkeit im Fahrkontext sei auf Wright, Stone, Horberry und Reed (2007) verwiesen. Im Folgenden findet sich eine Übersicht über die in den vorliegenden Studien verwendeten physiologischen Maße.

1.3.3.1 EEG-Alpha-Spindeln

Die Beschreibung und Evaluation der in unserer Arbeitsgruppe entwickelten Methodik zur automatisierten Erkennung und Parametrisierung von so genannten Alpha-Spindeln ist zentraler Inhalt von Studie 1. An dieser Stelle sollen lediglich kurz die Grundlagen sowie die erwarteten Vorteile des Maßes zusammengefasst werden. Die Basis stellt die Aktivität im EEG-Alpha-Band (7-13 Hz) dar, welche sich in diversen Studien im Realverkehr und in Simulatoren sensitiv für Müdigkeit gezeigt hat (Brookhuis & De Waard, 1993; Horne & Baulk, 2004; Lal & Craig, 2002; O'Hanlon & Kelly, 1977; Papadelis et al., 2007) und insbesondere in seiner parieto-okzipitalen Ausprägung im Zusammenhang mit visueller Reaktionsfähigkeit steht (Hanslmayr et al., 2007; van Dijk, Schoffelen, Oostenveld & Jensen, 2008). Im Gegensatz zur klassischen eindimensionalen Quantifizierung der Leistung werden bei der Analyse auf Grundlage der Alpha-Spindeln Informationen über das Auftreten einzelner Alpha-Ereignisse (Kecklund & Åkerstedt, 1993; Tietze & Hargutt, 2001), das heißt kurzzeitiger, monochromatischer Anstiege der Leistung im Alpha-Frequenzband, hinzugenommen. Dieses Vorgehen ermöglicht zum einen die Erfassung zusätzlicher Informationen wie Spindelrate (Auftretenshäufigkeit), -dauer, -amplitude und -frequenz und zum anderen eine geringere Anfälligkeit gegenüber Artefakten. Hierzu wird das Signal-Rausch-Verhältnis zum entsprechenden Zeitpunkt für die Identifikation einer Alpha-Spindel berücksichtigt. Weiterhin ist das monochromatische Spektrum einer Alpha-Spindel und die Verarbeitung allein dieser spektralen Eigenschaften weniger anfällig für breitbandiges Rauschen, welches sowohl von technischen als auch biologischen Artefakten stammen kann. Dies ist insbesondere im vorliegenden Anwendungskontext von Bedeutung.

1.3.3.2 P3-Amplitude des Ereigniskorrelierten Potentials

Ereigniskorrelierte Potentiale (EKPs) bieten die Möglichkeit, die neurophysiologische Verarbeitung eines Reizes mit hoher zeitlicher Auflösung im EEG abzubilden (für detaillierte Ausführungen siehe Luck, 2005; Rugg & Coles, 1993). Verschiedene Studien konnten zeigen, dass Amplitude und Latenz der P3-Komponente in Abhängigkeit des Vigilanzzustands variieren (Curcio et al., 2001; Gosselin, De Koninck & Campbell, 2005; Jackson et al., 2008; Koelega et al., 1992; für eine Übersichtsarbeit zur P3-Komponente siehe Polich, 2007). Da die Verarbeitung externer Informationen entscheidend für die Reaktionsfähigkeit ist, stellt diese Methode ein valides Messinstrument zur Erfassung des Fahrerzustands dar. Eingeschränkt wird ihr Einsatz jedoch dadurch, dass die Präsentation von Reizen eine notwendige Bedingung für die Auslösung des EKPs ist. Im Fall der P3 im Oddball-Paradigma muss sogar eine Vielzahl von Reizen dargeboten werden, die für die spätere Auswertung keine Relevanz hat, da die P3 sich nur auf einen im Vergleich seltenen Stimulus ausbildet. Duncan et al. (2009) geben eine Mindestanzahl von 36 Durchgängen für ein hinreichendes Signal-Rausch-Verhältnis und eine somit zuverlässigen Erfassung der P3 im Labor an, was die zeitliche Auflösung des Maßes weiter eingeschränkt. Somit bedarf es stets einer Abwägung zwischen dem Ziel, den Zustand durch wenige Reize möglichst wenig zu beeinflussen und dem Bestreben, durch die Präsentation vieler Reize einen möglichst exakten Indikator für die P3-Amplitude zu erhalten.

1.3.3.3 Herzrate

Die Herzrate hat sich im Kontext des Autofahrens wiederholt sensitiv gegenüber Müdigkeit gezeigt (Brookhuis & De Waard, 1993; Dureman & Bodén, 1972; O'Hanlon & Kelly, 1977). Sie stellt allerdings ein sehr globales Maß dar, welches in erster Linie die reduzierte periphere Aktivierung im Verlauf einer monotonen Fahraufgabe abbildet (Lal & Craig, 2002). Auf Grund der geringen Spezifität finden sich uneinheitliche Befunde bezüglich der Korrelation zwischen Herzrate und Müdigkeit (Hefner et al., 2009). Dies beruht auf dem sehr großen Einfluss müdigkeitsunspezifischer Ursachen wie zum Beispiel Muskelaktivität (Manzey, 1998). Ebenfalls zu berücksichtigen ist ein zirkadianer Einfluss auf die Herzrate. Während der Fahrtzeit der vorliegenden Erhebungen (12:30 - 16:30 Uhr) waren allerdings aus der Literatur keine bedeutsamen zirkadianen Effekte auf die Herzrate zu erwarten (Gander, Connell & Graeber 1986; Vandewalle et al., 2007).

1.3.3.4 Lidschlussdauer

Schleicher et al. (2008) untersuchten anhand einer umfangreichen Stichprobe (N = 129) eine Vielzahl auf Lidschlüssen, Sakkaden und Fixationen basierender okulomotorischer Parameter während einer monotonen Fahrsimulatoraufgabe. Die höchsten Korrelationen sowohl mit subjektiven als auch auf Grundlage von Verhaltensbeobachtungen bewerteten Müdigkeitsindikatoren ergaben sich für die Lidschlussdauer. Entsprechend identifizierten Caffier, Erdmann und Ullsperger (2003) aus diversen Lidschlussmaßen die Lidschlussdauer als den besten Parameter zur Erfassung von Müdigkeit und auch Ingre et al. (2006a) berichten einen starken Zusammenhang zwischen der Müdigkeitsbewertung anhand der KSS und der Lidschlussdauer in einer Simulatorstudie. Weiter fanden Campagne, Peybale und Muzet (2005) in einer monotonen Simulatorstudie eine Erhöhung der Lidschlussdauer mit zunehmender Fahrtzeit. Dabei wurde durch signifikante visuelle Ereignisse eine unmittelbare Reduktion der Lidschlussdauer ausgelöst, was auf eine gute zeitliche Auflösung und eine hohe Sensitivität für aktivierende Ereignisse schließen lässt. Auf Grund dieses Potentials entschieden wir uns für den zusätzlichen Einsatz der Lidschlussdauer zur Erfassung der kurzzeitigen Aktivierungseffekte in Studie 3.

1.3.4 Verhaltensbeobachtung

Obwohl die Verhaltensbeobachtung in den vorliegenden Studien keine Anwendung fand, kann auch sie eine valide Messmethode zur Erfassung von Müdigkeit darstellen. Ein häufig eingesetztes Ratingverfahren zur Beurteilung von Müdigkeit auf Grundlage einer Ansicht des Gesichts entwickelten Wierwille und Ellsworth (1994). Sie konnten zeigen, dass das Verfahren über verschiedene Rater reliable und konsistente Ergebnisse lieferte. Zudem wurden hohe Korrelationen mit diversen etablierten subjektiven und objektiven Müdigkeitsmaßen berichtet. Rogé und Muzet (2001) zeigten, dass während einer monotonen, simulierten Fahraufgabe die Anzahl aufgabenirrelevanter Bewegungen (z.B. selbstbezogene Gesten, Anpassung der Haltung) mit abnehmender Vigilanz signifikant zunahm.

1.4 Gegenmaßnamen zur Reduzierung von Müdigkeit beim Autofahren

Unabhängig von der Effektivität der Gegenmaßnahmen zur unmittelbaren situativen Reduzierung von Müdigkeit ist die beste langfristige und präventive Maßnahme eine angemessene Planung langer Autofahrten. Dies bedeutet, dass etwa in der Nacht vor der Fahrt genügend geschlafen und am Abend vorher auf Alkohol verzichtet werden sollte (Radun, 2009). Des Weiteren ist es ratsam, die Tageszeiten des größten Schlafbedürfnisses zu meiden. Dies widerspricht zwar paradoxerweise diversen Empfehlungen zur Entlastung der Straßen zur Hauptreisezeit, trägt aber zur Reduzierung des Unfallrisikos bei (Philip et al., 1996). Auch wenn durch diese edukativen Maßnahmen sicher ein bedeutsamer Anteil übermüdeter Fahrer auf den Straßen vermieden werden könnte, so scheint es doch unrealistisch, allein dadurch Müdigkeit von den Straßen verbannen zu können (MacLean et al., 2003; Radun, 2009). Daher ist eine Evaluation von unmittelbaren Maßnahmen während der Fahrt und in Fahrtpausen von größter Bedeutung.

May und Baldwin (2009) gehen davon aus, dass zwar die Erfassung von Müdigkeit unabhängig von ihrem Ursprung erfolgen kann, die Wirksamkeit verschiedener Gegenmaßnahmen aber entscheidend von dem eigentlichen Auslöser der Müdigkeit abhängt. In den meisten Studien werden Gegenmaßnahmen zu schlafbezogener Müdigkeit untersucht, so auch in den von Horne und Reyner veröffentlichten Arbeiten (Horne & Reyner, 1996; Reyner & Horne, 1997, 1998, 2000, 2002). Die Essenz ihrer Studien ist, dass schlafbezogene Müdigkeit kurzfristig und nachhaltig am effektivsten durch die Aufnahme von 150 mg Koffein (etwa eine große Tasse Kaffee) und einen anschließenden Schlaf von etwa 15 Minuten Dauer reduziert werden kann. Weitere untersuchte Maßnahmen wie etwa Frischluftzufuhr oder Radiohören zeigten lediglich kurzzeitige Effekte und wurden daher von den Autoren nicht zur nachhaltigen Anwendung empfohlen.

Royal (2003) befragte 4.010 Autofahrer danach, welche Maßnahmen sie selbst ergreifen, um Müdigkeit entgegenzuwirken. Die häufigsten Nennungen waren dabei das Pausieren oder Schlafen auf einem Rastplatz (43%), das Öffnen des Fensters (26%), der Konsum von koffein- und/oder zuckerhaltigen Getränken (17%) sowie das Einschalten des Radios mit hoher Lautstärke (14%). In Befragungen von Maycock (1997; 4621 LKW-Fahrer) sowie Nordbakke und Sagberg (2007; 1513 Autofahrer) wurde die empfundene Effektivität von Gegenmaßnahmen erfragt. Neben einem Fahrerwechsel wurden dabei folgende Maßnahmen am häufigsten genannt: Anhalten mit verschiedenen Aktivitäten bzw. Schlaf (Maycock et al., 1997: 57% / Nordbakke & Sagberg, 2007: 80%), Fensteröffnen (68% / 34%), Radiohören (30% / 13%); Gespräch mit dem Beifahrer (25% / 32%), Kaffeetrinken (14% / 22%), Gespräch am Mobiltelefon (- / 5%). Auch Häkkänen und Summala (2000) fanden, dass ca. 20% von 317 befragten LKW-Fahrern zur Aktivierung während der Fahrt mit dem Mobiltelefon telefonierten. Zusammenfassend legen diese Befunde nahe, dass zumindest subjektiv die verbale Interaktion beziehungsweise eine auditive Stimulation von vielen Fahrern in Phasen hoher Müdigkeit als hilfreiche Gegenmaßnahme empfunden wird. Ob dies auch objektiv der Fall ist, wurde im Rahmen von Studie 3 untersucht.

1.4.1 Effekte einer verbalen Interaktion während der Fahrt

Es existieren bislang kaum experimentelle Studien, die sich mit den Effekten einer verbalen Interaktion unter monotonen Fahrbedingungen befassen. Kürzlich untersuchten Chan und Atchley (2009) diese Fragestellung systematisch. Allerdings betrug die Fahrtdauer in der Simulatorstudie nur 45 Minuten und es wurden lediglich Fahrverhaltensmaße und Gedächtnisleistung als abhängige Variablen erfasst. So folgerten auch die Autoren, dass keine eindeutige Aussage über eine mögliche Aktivierung getroffen werden konnte.

Aktivierende Effekte unter monotonen Bedingungen könnten auch durch Nebenaufgaben mit einer verbalen Komponente und durch verbale Interaktionen mit einem Beifahrer oder am Mobiltelefon zustande kommen. So fand Drory (1985), dass die Leistung von LKW-Fahrern stieg, wenn sie von Zeit zu Zeit in einer verbalen Aufgabe involviert waren. Entsprechend berichten Oron-Gilad, Ronen und Shinar (2008) sowie Gershon, Ronen, Oron-Gilad und Shinar (2009), dass eine quizähnliche verbale Aufgabe eine Verschlechterung der Fahrleistungsmaße verhinderte und gleichzeitig physiologische Maße einen verbesserten Zustand indizierten. Negative Effekte in nicht-monotonen Situationen sowohl durch Konversationen am Mobiltelefon mit und ohne Freisprecheinrichtung als auch durch Unterhaltungen mit dem Beifahrer berichten hingegen Horrey und Wickens (2006) sowie Caird, Willness, Steel und Scialfa (2008) in ausführlichen Übersichtsarbeiten. McEvoy et al. (2005) leiteten aus australischen Unfalldaten eine vierfach erhöhte Wahrscheinlichkeit während der Nutzung eines Mobiltelefons zu verunfallen ab, unabhängig davon, ob dabei eine Freisprecheinrichtung genutzt wurde oder nicht. Folglich ist davon auszugehen, dass ein zu erwartender aktivierender und somit positiver Effekt auf die Verkehrssicherheit während der verbalen Interaktion möglicherweise durch eine gleichzeitige Ablenkung zunichte gemacht werden könnte.

Dies war der Grund, weshalb der Phase nach dem Ende einer verbalen Interaktion in Studie 3 aus der Anwendungssicht die größte Bedeutung beigemessen wurde. Als experimentelle Manipulation diente dabei die verbale Abfrage des subjektiven Fahrerzustandes. Der experimentelle Ansatz eignete sich dabei gleichzeitig für die in Abschnitt 1.3.1 erläuterte Ableitung einer angemessenen Abfragehäufigkeit ohne die nachhaltige Beeinflussung des zu erfassenden Müdigkeitszustands.

1.5 Fragestellungen

Die übergeordnete Zielstellung der vorliegenden Arbeit war die objektive und zuverlässige Erfassung des Müdigkeitszustands von Autofahrern im realen Straßenverkehr am Tage unter Verwendung von physiologischen Maßen sowie von Leistungsmaßen. Ein besonderes Augenmerk galt dabei der Zuverlässigkeit subjektiver Urteile sowie der Prüfung des Einflusses einer Abfrage auf den objektiven Zustand des Fahrers. Somit lassen sich drei zentrale Fragestellungen ableiten:

1) Eignen sich die auf EEG-Alpha-Spindeln basierenden Maße zur Erkennung eines Müdigkeitszustands während monotoner Tagfahrten? Haben diese Maße eine bessere Sensitivität und Spezifität als die etablierten, auf der Leistung in einem EEG-Frequenzband basierenden Maße?
2) Wie entwickeln sich subjektive, Verhaltens- und physiologische Maße im Verlauf einer monotonen Tagfahrt im Realverkehr? Stimmt der Verlauf der Selbsteinschätzung durch die Fahrer mit den objektiven Maßen überein?
3) Welchen Einfluss hat eine verbale Abfrage des Fahrerzustands während einer monotonen Tagfahrt auf die Fahrermüdigkeit, und wie lange hält ein möglicher aktivierender Effekt an?

2 Studien

Im Rahmen der Promotionsarbeiten wurden zwei umfangreiche Fahrversuche im Realverkehr durchgeführt, die im Folgenden als Erhebung 1 (2006) und Erhebung 2 (2007) bezeichnet werden. Für beide Erhebungen übernahm der Autor jeweils hauptverantwortlich die Gestaltung des experimentellen Designs, die Planung und Durchführung der Datenerhebung sowie die gruppenstatistische Auswertung sämtlicher präsentierter abhängiger Variablen. Die Stichprobengrößen sowie die Verwendung der Datensätze im Rahmen der drei vorgestellten Studien sind in Abbildung 2 dargestellt. Der prozentuale Anteil des Autors an den drei Studien lässt sich wie folgt beziffern:

Studie 1: Einleitung (75%); Methodik des Realfahrtteils (50%); Erhebung der Realdaten (80%); Auswertung der Realdaten (30%); Diskussion (30%)

Studie 2: Einleitung (75%); Methodik (75%); Datenerhebung (80%); Auswertung (80%); Diskussion (90%)

Studie 3: Einleitung (75%); Methodik (75%); Datenerhebung (80%); Auswertung (80%); Diskussion (90%)

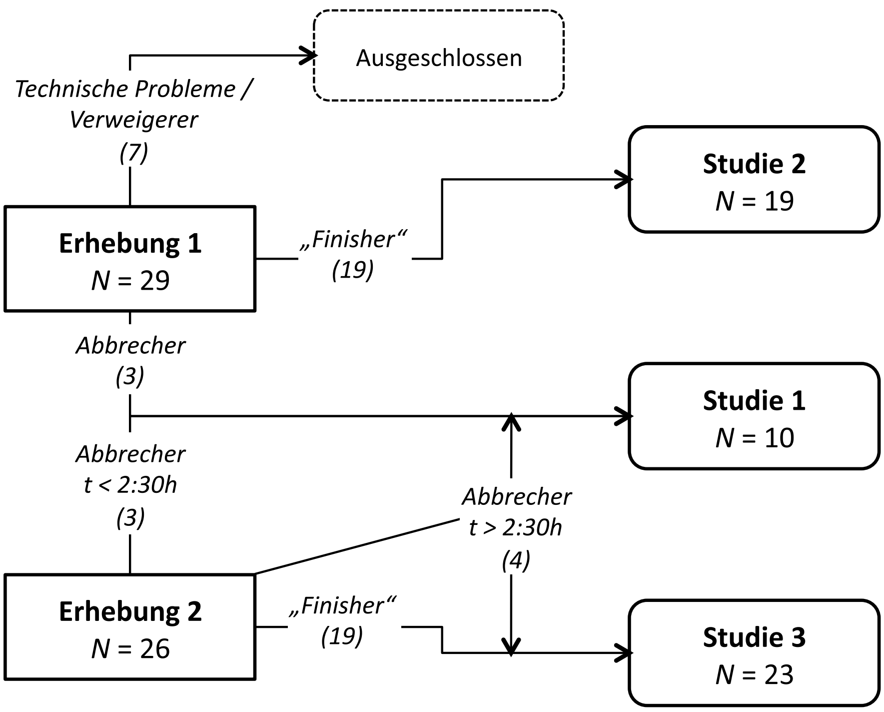

Abbildung 2: Verwendung der in Erhebungen 1 und 2 erfassten Datensätze im Rahmen der drei Studien.

2.1 Studie 1: EEG-Alpha-Spindelmaße als Indikatoren für Fahrermüdigkeit unter Realfahrtbedingungen

Zentrale Bestandteile der methodisch orientierten Studie 1 waren die Vorstellung der signalanalytischen Grundlagen des Alpha-Spindel-Detektors und seine Evaluation anhand simulierter sowie im Realverkehr aufgezeichneter Daten. Zunächst wurde der FFT-basierte (Fast Fourier Transformation) Algorithmus vorgestellt, der EEG-Alpha-Spindeln erkennt und deren charakteristische Eigenschaften Rate (Auftretenshäufigkeit pro Minute), Amplitude, Frequenz und Dauer bestimmt. Die Zuverlässigkeit und Genauigkeit des Algorithmus wurden anhand simulierter Daten, bei denen die tatsächliche Verteilung der Spindeln sowie die Eigenschaften der Spindeln bekannt waren, unter Beweis gestellt. In einem zweiten Schritt wurde der Algorithmus auf Datensätze angewendet, in denen die jeweiligen Fahrer eine monotone Tagfahrt im realen Straßenverkehr auf Grund von starker Müdigkeit abgebrochen hatten. Hierdurch konnte ein kritischer Müdigkeitsgrad definiert werden, ohne auf andere Referenzmaße und die damit verbundenen Unsicherheiten zurückgreifen zu müssen. Als unabhängige Variable wurde für 10 Probanden der Faktor Fahrtabschnitt in den Ausprägungen „müde“ (letzte 20 Minuten vor Fahrtabbruch) und „wach“ (erste 20 Minuten der Fahrt) definiert. Es konnte gezeigt werden, dass Rate, Dauer und Amplitude der Alpha-Spindeln höhere Effektstärken für den Faktor Fahrtabschnitt aufwiesen als die Leistung im Alpha-Band oder weitere auf der Leistung in anderen EEG-Bändern basierende Indikatoren (Delta, Theta, Beta, (Alpha+Theta)/Beta). Die spektrale Frequenz der Alpha-Spindeln zeigte hingegen keinerlei Korrelation mit dem Müdigkeitszustand. Zum Ausschluss möglicher, vom Müdigkeitslevel unabhängiger zeitlicher Effekte wurde aus Erhebungen 1 und 2 eine Vergleichsstichprobe (N = 31) von Nicht-Abbrechern herangezogen. Die Abbrecher zeigten gegenüber den Nicht-Abbrechern signifikant höhere Spindelraten und -dauern sowie einen größeren Effekt des Fahrtabschnitts, was für eine hohe Spezifität dieser Maße für Müdigkeit spricht. Erste Analysen deuten zudem auf eine gute Abbildbarkeit des Müdigkeitsverlaufs sowie auf eine zufriedenstellende zeitliche Auflösung hin. Zusammenfassend liefert die Alpha-Spindelrate somit den besten auf dem Spontan-EEG basierenden Indikator für die Erfassung von Müdigkeitszuständen im Straßenverkehr.

2.2 Studie 2: Fehleinschätzung des eigenen Vigilanzzustands durch Autofahrer während monotoner Tagfahrten

Zur Analyse der Effekte monotoner Tagfahrten auf den Vigilanzzustand des Fahrers sowie dessen Fähigkeit, seinen Zustand zu beurteilen, wurde ein Fahrversuch von 428 km Länge im realen Straßenverkehr durchgeführt. Neben der subjektiven Einschätzung des Zustands durch den Fahrer alle 20 Minuten auf den Dimensionen Müdigkeit, Aufmerksamkeit für die Fahraufgabe und Monotonie der Fahraufgabe wurden das Leistungsmaß langsame Reaktionszeiten in einem auditiven Oddball-

Paradigma sowie die physiologischen Maße Alpha-Spindelrate, P3-Amplitude des ereigniskorrelierten Potentials sowie Herzrate erhoben. Mit zunehmender Fahrtdauer indizierten sämtliche Maße eine kontinuierliche Reduzierung der Vigilanz. Im letzten Viertel der Fahrt berichteten die Fahrer subjektiv einen verbesserten Zustand, während im Gegensatz dazu die objektiven Maße allesamt in Richtung einer weiteren Verschlechterung des Zustands deuteten. Dieser Befund belegt eine mangelnde Selbsteinschätzung der Fahrer nach etwa 3 Stunden monotonen Autofahrens. Es ist nicht möglich eindeutig zu folgern, ob diese Fehleinschätzung durch die verbesserte zirkadiane Phase, die erhöhte Verkehrsdichte oder die Erwartung des Fahrtendes bedingt wurde. Die Befunde unterstützen zum einen die Wichtigkeit edukativer Maßnahmen mit dem Ziel einer Aufklärung über die Grenzen menschlicher Selbsteinschätzung in Bezug auf Müdigkeitsphänomene beim Autofahren. Zum anderen wird die Notwendigkeit der Entwicklung objektiver Systeme zur zuverlässigen Erfassung von Müdigkeit im Straßenverkehr belegt.

2.3 Studie 3: Der kurzzeitige Effekt einer verbalen Abfrage des Fahrerzustands auf Vigilanzindikatoren während monotoner Tagfahrten

In einem monotonen Tagfahrversuch wurden die objektiven Vigilanzmaße EEG-Alpha-Spindelrate, Lidschlussdauer, Herzrate, P3-Amplitude des ereigniskorrelierten Potentials sowie Reaktionszeiten erfasst und in zeitlicher Beziehung zu verbalen Abfragen des subjektiven Fahrerzustands ausgewertet. Die Abfragen fanden alle 20 Minuten statt und dauerten etwa 60 Sekunden. Der in Studie 2 gefundene Effekt der Fahrtdauer konnte für alle Maße bis auf die P3-Amplitude repliziert werden. Es zeigte sich zudem, dass die kontinuierlichen physiologischen Maße (Alpha-Spindelrate, Lidschlussdauer, Herzrate) eine signifikante Verbesserung des Vigilanzzustands während und nach der Abfrage im Vergleich zum Ausgangszustand vor der Abfrage indizierten. Nach Ende der Abfrage hielt der aktivierende Effekt dabei für maximal zwei Minuten an. Die langsamen und entgegen der Erwartung ebenso die schnellen Reaktionszeiten zeigten der Abfrage folgend eine Verbesserung. Die P3-Amplitude zeigte auch hier keine signifikante Veränderung. Die Verbesserung auch der schnellen Reaktionszeiten lässt sich am ehesten durch einen motivationalen Einfluss der Interaktion mit dem Versuchsleiter erklären. Das Ausbleiben jeglicher Effekte der P3-Amplitude ist vermutlich bedingt durch eine für den angewandten Kontext zu geringe Anzahl von Stimuli je ausgewerteter Zeiteinheit. Aus den Befunden lässt sich schlussfolgern, dass eine verbale Abfrage einen deutlich aktivierenden Einfluss auf den Fahrerzustand hat, auch wenn der Effekt nur von kurzer Dauer ist. Folglich ist davon auszugehen, dass bis zu einer Häufigkeit von einer Abfrage etwa alle fünf Minuten ein Großteil der Daten unbeeinflusst von Abfrageeffekten verbleiben. Eine verbale Interaktion zeigt somit nur ein bedingtes Potential für einen Einsatz als Maßnahme gegen Fahrermüdigkeit. Es muss jedoch davon ausgegangen werden, dass ein aktivierender Effekt nur von kurzer Dauer ist und der Fahrer sehr schnell wieder in seinen vor der Interaktion vorliegenden Zustand zurückfällt.

3 Allgemeine Diskussion und Ausblick

Zusammengefasst konnte in den vorgestellten Studien gezeigt werden, dass

(a) die auf EEG-Alpha-Spindeln basierenden Maße Spindelrate, Spindeldauer und Spindelamplitude gegenüber klassischen EEG-Leistungsmaßen eine höhere Sensitivität und Spezifität für die Erkennung von Müdigkeit im Kontext des monotonen Autofahrens zeigen;

(b) eine von Autofahrern subjektiv empfundene Verbesserung ihres Müdigkeitszustands zum Ende einer vierstündigen monotonen Tagfahrt im Gegensatz zu der weiterhin fortschreitenden Verschlechterung objektiver Müdigkeitsindikatoren steht;

(c) der Prozess der verbalen Abfrage des subjektiven Zustands während einer monotonen Tagfahrt in einer unmittelbaren, signifikanten Aktivierung resultiert, welche nach Ende der Abfrage lediglich für maximal zwei Minuten anhält;

Diese Befunde belegen die Nützlichkeit von Alpha-Spindelmaßen zur EEG-basierten Erfassung von Müdigkeit und stellen gleichzeitig die Zuverlässigkeit von Selbsteinschätzungen in Frage. Zudem wird erstmals der unmittelbar aktivierende Einfluss einer verbalen Abfrage des Fahrerzustands mit einer hohen zeitlichen Auflösung aufgezeigt. Trotz Klärung der eingangs formulierten Fragestellungen zeigen die Befunde auch eine Reihe weiterer Ansatzpunkte für zukünftige Forschungen auf, die im Folgenden diskutiert werden.

Ein möglicher Kritikpunkt am methodischen Vorgehen in Studie 1 findet sich in der alleinigen Definition der unabhängigen Variable an Hand des Abbruchkriteriums. Dies erscheint jedoch als die Option mit höchster ökologischer Validität, da sich die Müdigkeit nicht nur in einem hohen KSS-Wert, sondern insbesondere in der verhaltensrelevanten Entscheidung abzubrechen widerspiegelt. Zudem waren die Versuchsleiter auf Grund ihrer Beobachtungen in allen zehn Fällen davon überzeugt, dass tatsächlich Müdigkeit den entscheidenden Auslöser für den Abbruch der Fahrt darstellte. Dennoch wird es für weitere Studien interessant sein, die Alpha-Spindelmaße an weiteren Außenkriterien zu validieren. So ist zum Beispiel noch nicht hinreichend geklärt, welches die plausibelsten Verhaltenskorrelate der Alpha-Spindelmaße sind. Ein potentieller Kandidat ist dabei die visuelle Reaktionsfähigkeit, da für diese bereits Zusammenhänge mit dem Alpha-Band bestätigt werden konnten. Ebenso informativ wie die gemittelte Analyse einer Vielzahl von Spindeln könnte dabei auch die Charakterisierung einzelner Spindeln sein. Auch wird es wichtig sein, genauer zu beleuchten, was den Zustand beim Auftreten einer Spindel mit einer bestimmten Charakterisierung von spindelfreien Phasen oder Phasen mit Spindeln anderer Charakteristika unterscheidet. Sollten zu dieser Frage Fortschritte erzielt werden, so hätte dies auch bedeutsame Implikationen für die zeitliche Auflösung der Spindelmaße.

Die zeitliche Auflösung der Spindelmaße sowie ihre Güte in Bezug auf die Erfassung von Müdigkeitsverläufen sind zwei eng zusammenhängende Aspekte, die in Studie 1 zwar adressiert wurden, im Weiteren allerdings einer detaillierteren Betrachtung bedürfen. Die Befunde aus den Studien 1 und 2 deuten auf eine zuverlässige Erfassung auch solcher Zustände hin, die auf dem Müdigkeitskontinuum zwischen dem wachen Ausgangszustand und dem müden Zustand vor Abbruch der Fahrt liegen. In Bezug auf die zeitliche Auflösung wurde in einem ersten Ansatz in Studie 1 die Trennung der zwei definierten Zustände auf Grundlage von Zeitfenstern unterschiedlicher Dauer getestet. Damit gelang es zwar abzuleiten, dass auf Grund der geringen Ereignisrate der Spindeln die Länge der Zeitfenster nicht kleiner als drei Minuten gewählt werden sollte, allerdings konnten durch das Fehlen einer zeitlich hoch aufgelösten „ground truth" innerhalb des als müde definierten Zeitfensters keine eindeutige Aussage über die untere Grenze der zeitlichen Auflösung getroffen werden. Zur endgültigen Klärung beider Fragestellungen bedarf es einer Studie, in der sowohl die Zwischenzustände als auch moderate Variationen des Zustands mit einer hohen zeitlichen Auflösung vorzugsweise über Leistungskriterien in einer kontinuierlichen Aufgabe definiert sind.

Auf Grund der hohen interindividuellen Variabilität sowohl im Grundniveau der Spindelmaße als auch in ihrer räumlichen Topographie (siehe Studie 1, Abbildung 7) ist davon auszugehen, dass eine individuelle Detektion von Müdigkeit eine individualisierte Auswertung erfordert. Von daher ist insbesondere unter Gesichtspunkten der Anwendung die Eignung der Alpha-Spindelmaße für die diagnostische Erkennung von Müdigkeitszuständen im Sinne einer Klassifikation für einzelne Individuen von großer Bedeutung. Eine interessante Fragestellung bleibt dabei, ob ein Klassifikationsalgorithmus für diese Adaptation die Kenntnis zweier klar definierter Zustände von Wachheit und Müdigkeit zwingend erfordert, oder ob es im Sinne einer effizienten Lernphase auch gelingen kann, auf Grundlage eines bekannten beziehungsweise eines in der Trainingsphase nur sehr gering variierenden Zustands den Klassifikator zu implementieren.

Studie 2 liefert eindeutige Hinweise für eine verschlechterte Selbsteinschätzungsfähigkeit beim monotonen Autofahren am Tage. Dabei mag zwar kritisch argumentiert werden, dass die Fahrer sich durchschnittlich im mittleren Müdigkeitsniveau bewegten und damit die Befunde im Vergleich zu Effekten stärkerer Müdigkeitsstadien von geringerer praktischer Relevanz sind. Dem ist allerdings entgegenzuhalten, dass auch und gerade in frühen Müdigkeitsstadien der Selbsteinsicht eine große Bedeutung zukommt, um etwa frühzeitig eine Pause einzulegen und damit Effekten passiver aufgabenbezogener Müdigkeit rechtzeitig entgegenzuwirken, die sich auch in der vorliegenden Studie in einer signifikanten Reduzierung der Reaktionsfähigkeit manifestierten.

Eine aus den Befunden in Studie 2 nicht endgültig zu klärende Frage stellt der Auslöser der Fehleinschätzung dar. Deshalb sollten in Folgestudien die betreffenden Faktoren - zirkadiane Phase, Erwartung des Fahrtendes, erhöhte Verkehrsdichte - unabhängig voneinander variiert werden. Die Identifikation des zentralen Auslösers könnte in einer präziseren Ableitung von Empfehlungen aus den vorliegenden Ergebnissen resultieren, aber auch in fahrzeugbasierten Algorithmen zur Müdigkeitserkennung und zugehörigen Warnstrategien Anwendung finden.

Entsprechend könnten für Studie 3 Erkenntnisse über die Generalisierbarkeit des aktivierenden Effekts auf andere Fahrtkontexte zu einer besseren Anwendbarkeit der Ergebnisse beitragen. Dafür wäre es nötig herauszufinden, unter welchen Umgebungsbedingungen, zu welcher Tageszeit und insbesondere bei welcher Art von Müdigkeit es zu einer Aktivierung durch eine verbale Interaktion kommt und vor allem ob und unter welchen Bedingungen in der Nettobetrachtung (Aktivierung vs. Ablenkung) ein positiver Einfluss auf die Verkehrssicherheit zu erwarten ist.

Wie in der Einleitung beschrieben, wird die Validität von Reaktionszeiten zur Erfassung von Fahrermüdigkeit kontrovers diskutiert. Die in Studien 2 und 3 beobachteten Übereinstimmungen der Befunde der langsamen Reaktionszeiten mit denen der restlichen objektiven Maße deuten auf eine gute Validität der langsamen Reaktionszeiten hin. In beiden Studien zeigten Mittelwert- beziehungsweise Medianmaße, die in den Artikeln der Übersichtlichkeit halber nicht berichtet wurden, keine oder deutlich geringere Effekte als die langsamen Reaktionszeiten. Als unbefriedigend im Sinne eines Einsatzes als alleiniges Müdigkeitsmaß ist allerdings die zeitliche Auflösung der Reaktionszeitmaße zu bewerten, insbesondere weil die isolierte Auswertung der langsamsten Reaktionszeiten die Anzahl der verwendbaren Reaktionen deutlich reduziert. Gerade im angewandten Kontext ist nicht davon auszugehen, dass einzelne oder nur sehr wenige Reaktionszeiten sich als brauchbare Indikatoren für Müdigkeit erweisen, da der unabhängig von Müdigkeitseffekten auftretende Varianzanteil auf Grundlage sehr weniger Reaktionen stets den müdigkeitsbedingten Anteil übertreffen und ein plausibles Muster für Müdigkeitseffekte sich erst nach Analyse einer hinreichenden Anzahl von Reaktionen abbilden dürfte. Dies war Veranlassung dafür, nach den Erfahrungen aus Erhebung 1 die Antworthäufigkeit in Erhebung 2 zu erhöhen und die Probanden im Oddball-Paradigma sowohl auf häufige als auch auf seltene Töne reagieren zu lassen. Die Tatsache, dass in beiden Studien im Mittel ein ähnliches Müdigkeitsniveau berichtet wurde und in Erhebung 2 tendenziell sogar mehr Probanden die Fahrt vorzeitig abbrachen, spricht dabei gegen die Befürchtung, dass eine häufigere Antwortrate eine Reduktion der Monotonie und somit des Müdigkeitszustands erzeugt.

Durch die Konstruktion der Zweitaufgabe als Oddball-Paradigma war es in den vorgestellten Erhebungen möglich, gleichzeitig die vigilanzsensitive P3-Amplitude im ereigniskorrelierten Potential zu erfassen. Die Erfahrungen aus den Studien 2 und 3 geben allerdings Anlass, einen kritischen Blick auf die Anzahl der in die Analyse eingehenden Durchgänge zu werfen. So gingen in Studie 2 etwa 100 Durchgänge je Fahrtabschnitt in die Auswertung ein, wodurch ein hinreichendes Signal-Rausch-Verhältnis erreicht wurde und sich ein Effekt der Fahrtdauer deutlich abbildete. In Studie 3 hingegen konnten auf Grund der geringeren Länge der ausgewerteten Zeitfenster im Mittel nur weniger als die 36 für Laborbedingungen geforderten Durchgängen (Duncan et al., 2009) ausgewertet werden. Folglich ist der Einsatz der P3-Amplitude nur im Rahmen von größeren Zeitabschnitten mit einem relativ stabilen Zustand zu empfehlen, in denen eine hinreichende Anzahl von auswertbaren Durchgängen akquiriert werden kann.

Betrachtet man die Ergebnisse aus den Studien 2 und 3, so ist es insbesondere die Konsistenz der Ergebnisse über verschiedene objektive Messmethoden hinweg, die im Sinne einer konvergenten Validität (Moller et al., 2006) für eine hohe Zuverlässigkeit der getroffenen Aussagen spricht. Aus dieser Perspektive ist es auch für weitere Studien zu empfehlen, eine Vielzahl von Maßen zu erfassen. Auf längere Sicht ist es allerdings insbesondere aus Gründen der Effizienz erstrebenswert, zu einem „Goldstandard", also einem einzelnen hoch validen Messinstrument zu gelangen. Die in der vorliegenden Arbeit präsentierte Alpha-Spindelrate und auch die Alpha-Spindeldauer könnten das Potential haben, diesen Goldstandard zu definieren. Um diesen Nachweis zu erbringen und vor allem die Bedingungen zu definieren, in denen dies der Fall ist, sind allerdings weitergehende Forschungsarbeiten nötig.

Zusammenfassend legen die Befunde aus allen drei Studien nahe, dass dem Thema Müdigkeit im Straßenverkehr weiter ein hohes Maß an Aufmerksamkeit entgegengebracht werden sollte und dass trotz stetiger Fortschritte im Erkenntnisstand weiterhin wichtige Forschungsfragen unbeantwortet bleiben. Die relative Wichtigkeit des Themas wird vermutlich in den kommenden Jahren durch gesellschaftliche und technische Entwicklungen befördert werden. Die zunehmende Automatisierung in der Fahrzeugführung sowie die fortschreitende Entwicklung der Gesellschaft zu einer 24-Stunden Gesellschaft lassen annehmen, dass sich ohne entsprechende Interventionen höhere Fallzahlen von sowohl passiver aufgabenbezogener Müdigkeit also auch schlafbezogener Müdigkeit im Straßenverkehr zu erwarten sind. Auch eine vollkommene Automatisierung des Straßenverkehrs (autonomes Fahren) löst die Problematik vorerst nicht auf, sondern rückt sie vielmehr weiter in den Mittelpunkt. So ist es allein aus rechtlichen Gründen unerlässlich, dass der Fahrer die Fahrsituation aufmerksam verfolgt und somit im der Regelkreis zwischen Fahrer und Fahrzeug aufrecht erhalten wird, da er stets verantwortlich in das Fahrgeschehen eingreifen können muss. Die Kunst wird somit darin bestehen, den Fahrer in einem Zustand zu halten, in dem er schnell und zuverlässig reagieren kann, obwohl durch die sinkenden Anforderungen der monotone Charakter der Fahraufgabe verstärkt wird.

Literatur

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience, 52*, 29-37.

Åkerstedt, T. (2000). Consensus statement: fatigue and accidents in transport operations. *Journal of Sleep Research, 9*, 395-395.

Åkerstedt, T., Kecklund, G., & Axelsson, J. (2008). Effects of context on sleepiness self-ratings during repeated partial sleep deprivation. *Chronobiology International, 25(2&3)*, 271-278.

Arnedt, J. T., Ainsley, M., Geddes, C., & MacLean, A. W. (2005). Comparative sensitivity of a simulated driving task to self-report, physiological, and other performance measures during prolonged wakefulness. *Journal of Psychosomatic Research, 58*, 61-71.

Arnedt, J. T., Wilde, G. J. S., Munt, P. W., & MacLean, A. W. (2001). How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? *Accident Analysis and Prevention, 33*, 337-344.

Baranski, J. V. (2007). Fatigue, sleep loss, and confidence in judgment. *Journal of Experimental Psychology: Applied, 13(4)*, 182-196.

Baulk, S. D., Reyner, L. A., & Horne, J. A. (2001). Driver sleepiness: evaluation of reaction time measurement as a secondary task. *Sleep, 24(6)*, 695-698.

Belz, S. M., Robinson, G. S., & Casali, J. G. (2004). Temporal separation and self rating of alertness as indicators of driver fatigue in commercial motor vehicle operators. *Human Factors, 46(1)*, 154-169.

Biggs, S. N., Smith, A., Dorrian, J., Reid, K., Dawson, D., van den Heuvel, C., & Baulk, S. (2007). Perception of simulated driving performance after sleep restriction and caffeine. *Journal of Psychosomatic Research, 63(6)*, 573-577.

Brookhuis, K., & De Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics, 36(9)*, 1099-1110.

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

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology, 89*, 319-325.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis and Prevention, 40*, 1282-1293.

Campagne, A., Peybale, T., & Muzet, A. (2005). Oculomotor changes due to road events during prolonged monotonous simulated driving. *Biological Psychology, 68*, 353-386.

CARE - EU road accidents database (2009). *Road safety evolution in EU*. Abgerufen am 07.02.2010 von http://ec.europa.eu/transport/road_safety/observatory/doc/historical_evol.pdf

Carskadon, M. A., Dement, W. C., Mitler, M. M., Roth, T., Westbrook, P. R., & Keenan, S. (1986). Guidelines for the multiple sleep latency test (MSLT) - a standard measure of sleepiness. *Sleep, 9*, 519-524.

Chan, M., & Atchley, P. (2009). Effects of cell phone conversations on driver performance while driving under highway monotony. *Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, MT*, 140-146.

Curcio, G., Casagrande, M., & Bertini, M. (2001). Sleepiness: evaluating and quantifying methods. *International Journal of Psychophysiology, 41*, 251-263.

Dawson, D., & Reid, K. (1997). Fatigue, alcohol and performance impairment. *Nature, 388*, 235-235.

Desmond, P. A., & Hancock, P. A. (2001). Active and passive fatigue states. In P. A. Hancock, & P. A. Desmond (Eds.), *Stress workload and fatigue (pp. 455-465)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Desmond, P.A., & Matthews, G. (2009). Individual differences in stress and fatigue in two field studies of driving. *Transportation Research Part F, 12*, 265-276.

DESTATIS (2009). *Unfallentwicklung auf Deutschen Straßen 2008*. Statistisches Bundesamt, Wiesbaden.

Dijk, D.-J., & von Schantz, M. (2005). Timing and consolidation of human sleep, wakefulness, and performance by a symphony of oscilators. *Journal of Biological Rhythms, 20*, 279-290.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers, 17*, 625-655.

Drory, A. (1985). Effects of rest and secondary task on simulated truck-driving task performance. *Human Factors, 27(2)*, 201-207.

Duncan, C., Barry, R., Connolly, J., Fischer, C., Michie, P., Näätänen, R., Polich, J., Reinvang, I., & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology, 120(11)*, 1883-1908.

Dureman, E. I., & Bodén, C. (1972). Fatigue in simulated car driving. *Ergonomics, 15(3)*, 299-308.

Eriksen, C. A., Åkerstedt, T., Kecklund, G., & Åkerstedt, A. (2005). Comment on short-term variation in subjective sleepiness. *Perceptual and Motor Skills, 101(3)*, 943-948.

Fairclough, S. H., & Venables, L. (2006). Prediction of subjective states from psychophysiology: a multivariate approach. *Biological Psychology, 71*, 100-110.

Folkard, S. (1997). Black times: temporal determinants of transport safety. *Accident Analysis and Prevention, 29(4)*, 417-430.

Gander, P. H., Connell, L. J., & Graeber, R. C. (1986). Masking of the circadian rhythms of heart rate and core temperature by rest-activity cycle in man. *Journal of Biological Rhythms, 1(2)*, 119-135.

George, C. F. P. (2003). Driving simulators in clinical practice. *Sleep Medicine Reviews, 7*, 311-320.

George, C. F., Nickerson, P., Hanly, P., Millar, T., & Kryger, M. (1987). Sleep apnea patients have more automobile accidents. *Lancet, 330(8556)*, 447.

Gershon, P., Ronen, A., Oron-Gilad, T., & Shinar, D. (2009). The effects of an interactive cognitive task (ICT) in suppressing fatigue symptoms in driving. *Transportation Research Part F, 12*, 21-28.

Gosselin, A., De Koninck, J., & Campbell, K. B. (2005). Total sleep deprivation and novelty processing: implications for frontal lobe functioning. *Clinical Neurophysiology, 116*, 211-222.

Graw, P., Kräuchi, K., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2004). Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. *Physiology & Behavior, 80*, 695-701.

Häkkänen, H., & Summala, H. (2000). Sleepiness at work among commercial truck drivers. *Sleep, 23*, 49-57.

Hancock, P. A., & Verwey, W. B. (1997). Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention, 29*, 495-506.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage, 37*, 1465-1473.

Hefner, R., Edwards, D., Heinze, C., Sommer, D., Golz, M., Sirois, B., & Trutschel, U. (2009). Operator fatigue estimation using heart rate measures. *Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, MT*, 110-117.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: a new approach. *Psychophysiology, 10*, 431-436.

Horne, J. A., & Baulk, S. D. (2004). Awareness of sleepiness when driving. *Psychophysiology, 4*, 97-110.

Horne, J. A., & Reyner, L. A. (1996). Counteracting driver sleepiness: Effects of napping, caffeine, and placebo. *Psychophysiology, 33*, 306-309.

Horne, J. A., & Reyner, L. A. (1999). Vehicle accidents related to sleep: a review. *Occupational and Environmental Medicine, 56*, 289-294.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors, 48*, 196-205.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006a). Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *Journal of Sleep Research, 15*, 47-53.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., Kecklund, G., & Pickles, A. (2006b). Subjective sleepiness and accident risk avoiding the ecological fallacy. *Journal of Sleep Research, 15*, 142-148.

Iudice, A., Bonanni, E., Gelli, A., Frittelli, C., Iudice, G., Cignoni, F., Ghicopulos, I., & Murri, L. (2005). Effects of prolonged wakefulness combined with alcohol and hands-free cell phone divided attention tasks on simulated driving. *Human Psychopharmacology, 20*, 125-132.

Jackson, M. L., Croft, R. J., Owens, K., Pierce, R. J., Kennedy, G. A., Crewther, D., & Howard, M. E. (2008). The effect of acute sleep deprivation on visual evoked potentials in professional drivers. *Sleep, 31(9)*, 1261-1269.

Johns M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep, 14*, 540-545.

Johns, M. W. (2000). A sleep physiologist's view of the drowsy driver. *Transportation Research Part F, 3*, 241-249.

Kaida, K., Åkerstedt, T., Kecklund, G., Nilsson, J. P., & Axelsson, J. (2007). The effects of asking for verbal ratings of sleepiness on sleepiness and its masking effects on performance. *Clinical Neurophysiology, 118*, 1324-1331.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology, 117*, 1574-1581.

Kecklund, G., & Åkerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics, 36*, 1007-1017.

Kecklund, G., Åkerstedt, T., Sandberg, D., Wahde, M., Dukic, T., Anund, A., & Hjälmdahl, M. (2006). State of the art review of driver sleepiness. DROWSI Project Report, Deliverable 1.1. Abgerufen am 03.01.2010 von http://www.vti.se/11934.epibrw.

Koelega, H. S., Verbaten, M. N., van Leeuwen, T. H., Kenemans, J. L., Kemner, C., & Sjouw, W. (1992). Time effects on event-related brain potentials and vigilance performance. *Biological Psychology, 34*, 59-86.

Lal, S. K. L., & Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological Psychology, 55*, 173-194.

Lal, S. K. L., & Craig A. (2002). Driver fatigue: electrophysiology and psychological assessment. *Psychophysiology, 39*, 1-9.

Lal, S. K. L., & Craig, A. (2005). Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *International Journal of Psychophysiology, 55*, 137-143.

Laurell, H., & Lisper, H.-O. (1978). A validation of subsidiary reaction time against detection of roadside obstacles during prolonged driving. *Ergonomics, 21*, 81-88.

Lenné, M. G., Triggs, T. J., & Redman, J. R. (1997). Time of day variations in driving performance. *Accident Analysis and Prevention, 29(4)*, 431-437.

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences, 1129*, 305-322.

Lin, C. T., Wu, R. C., Jung, T. P., Liang, S. F., & Huang, T. Y. (2005). Estimating driving performance based on EEG spectrum analysis. *EURASIP Journal on Applied Signal Processing, 19*, 3165-3174.

Lisper, H.-O., Laurell, H., & van Loon, J. (1986). Relation between time to falling asleep behind the wheel on a closed track and changes in subsidiary reaction time during prolonged driving on a motorway. *Ergonomics, 29(3)*, 445-453.

Liu, C. C., Hosking, S. G., & Lenné, M. G. (2009). Predicting driver drowsiness using vehicle measures: recent insights and future challenges. *Journal of Safety Research, 40*, 239-245.

Luck, S. J. (2005). *An introduction to the event -related potential technique*. Cambridge, USA: The MIT Press.

Mackworth, N. H. (1957). Vigilance. *The Advancement of Science, 53*, 389-393.

MacLean, A. W., David, D. R. T., & Thiele, K. (2003). The hazards and prevention of driving while sleepy. *Sleep Medicine Reviews, 7(6)*, 507-521.

Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Roesler (Ed.), *Ergebnisse und Anwendungen der Psychophysiologie. Enzyklopädie der Psychologie: Vol. C/I/5. Biologische Psychologie (pp. 799 -864)*. Göttingen: Hogrefe.

Matsumoto, Y., Mishima, K., Satoh, K., Shimizu, T., & Hishikawa, Y. (2002). Physical activity increases the dissociation between subjective sleepiness and objective performance levels during extended wakefulness in human. *Neuroscience Letters, 326*, 133-136.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasures technologies. *Transportation Research Part F, 12*, 218-224.

Maycock, G. (1997). Sleepiness and driving: the experience of U.K. car drivers. *Accident Analysis and Prevention, 29*, 453-462.

McDonald, N. (1984). *Fatigue, Safety and the Truck Driver*. London: Taylor and Francis.

McEvoy, S. P., Stevenson, M. R., McCartt, A. T., Woodwart, M., Haworth, C., Palamara, P., & Cercarelli, R. (2005). Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. *British Medical Journal*, doi:10.1136/bmj.38537.397512.55 (published 12 July 2005).

Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans & R. Schwing (Eds.), *Human Behavior and Traffic Safety (pp. 485-520)*. New York: Plenum Press.

Mitler, M. M., Gujavarty, K. S., & Browman, C. P. (1982). Maintenance of wakefulness test: a polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. *Electroencephalography and Clinical Neurophysiology, 53*, 658-661.

Moller, H. J., Kayumov, L., Bulmash, E. L., Nhan J., & Shapiro C. M. (2006). Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness. *Journal of Psychosomatic Research, 61*, 335-342.

Monk, T. (1987). Subjective ratings of sleepiness - the underlying circadian mechanisms. *Sleep, 10(4)*, 343-353.

Murray, E. J. (1965). *Sleep, Dreams, and Arousal*. New York: Appleton-Century-Crofts.

Nordbakke, S., & Sagberg, F. (2007). Sleepy at the wheel: Knowledge, symptoms and behaviour among car drivers. *Transportation Research Part F, 10*, 1-10.

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.

O'Hanlon J. F., & Kelly G. R. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In R. Mackie (Ed.), *Vigilance (pp. 87-109)*. New York: Plenum Press.

Oken, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology, 117*, 1885-1901.

Oron-Gilad, T., Ronen, A., & Shinar, D. (2008). Alertness maintaining tasks while driving. *Accident Analysis and Prevention, 40*, 851-860.

Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., & Schwab, C. W. (1995). Characteristics of crashes attributed to the driver having fallen asleep. *Accident Analysis and Prevention, 27*, 769-775.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E., & Maglaveras, N. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology, 118*, 1906-1922.

Parasuraman R. (1998). Brain system of vigilance. In R. Parasuraman (Ed.), *The attentive brain (pp. 221-256)*. Cambridge, MA: The MIT Press.

Philip, P., Ghorayeb, I., Leger, D., Menny, J. C., Bioulac, B., Dabadie,. P., & Guilleminault, C. (1997). Objective measurement of sleepiness in summer vacation long-distance drivers. *EEG Journal of Clinical Neurophysiolgy, 102*, 383-389.

Philip, P., Ghorayeb, I., Stoohs, R., Menny, J. C., Dabadie, P., Bioulac, B., & Guilleminault, C. (1996). Determinants of sleepiness in automobile drivers. *Journal of Psychosomatic Research, 41*, 279-288.

Philip, P., Sagaspe, P., Taillard, J., Guilleminault, C., Sanchez-Ortuno, M., Åkerstedt, T., & Bioulac, B. (2003). Fatigue, sleep restriction, and performance in automobile drivers: a controlled study in a natural environment. *Sleep, 26(3)*, 277-280.

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Åkerstedt, T., Charles, A., & Bioulac, B. (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep, 28(12)*, 1511-1516.

Phipps-Nelson, J., Redman, J. R., Schlangen, L. J. M., & Rajaratnam, S. M. W. (2009). Blue light exposure reduces objective measures of sleepiness during prolonged nighttime performance testing. *Chronobiology International, 26(5)*, 891-912.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118(10)*, 2128-2148.

Popp, R. F. J. (2005). *Gegenmaßnahmen bei Schläfrigkeit: Der Effekt von kurzwelligem Licht und olfaktorischer Stimulation*. Inaugural-Dissertation, Universität Regensburg.

Radun, I. (2009). *Fatigued driving: prevalence, risk factors and groups, and the law*. Dissertation. Helsinki: University Printing House.

Reyner, L. A., & Horne, J. A. (1997). Suppression of sleepiness in drivers: combination of caffeine with a short nap. *Psychophysiology, 34*, 721-725.

Reyner, L. A., & Horne, J. A. (1998). Evaluation of ‘‘in car’’ countermeasures to driver sleepiness: cold air and radio. *Sleep, 21*, 46-50.

Reyner, L. A., & Horne, J. A. (2000). Early morning driver sleepiness: effectiveness of 200 mg caffeine. *Psychophysiology, 37*, 251-256.

Reyner, L. A., & Horne, J. A. (2002). Efficacy of a 'functional energy drink' in counteracting driver sleepiness. *Physiology and Behavior, 75*, 331-335.

Rogé, J., & Muzet, A. (2001). Variations of the level of vigilance and of behavioural activities during simulated automobile driving. *Accident Analysis and Prevention, 33*, 181-186.

Royal, D. (2003). *National survey of distracted and drowsy driving attitudes and behavior. Volume 1: Findings (NHTSA, DOT HS 809 566).* Washington, DC: US Department of Transportation.

Rugg, M. D., & Coles, M. G. H. (1997). *Electrophysiology of mind. Event - related brain potentials and cognition*. Oxford: Oxford University Press.

Sagberg, F. (1999). Road accidents caused by drivers falling asleep. *Accident Analysis and Prevention, 31*, 639-649.

Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics, 51(7)*, 982-1010.

Smith, M., Oppenhuis, M., & Koorey, G. (2006). Fatigue crashes: the extent to which terrain change has an influence on the fatigued (drowsy) driver. *Proceedings of 22nd ARRB Conference - Research into Practice, Canberra, Australia (pp. 1-19)*.

Tejero Gimeno, P., Pastor Cerezuela, G., & Chóliz Montañés, M. (2006). On the concept and measurement of driver drowsiness, fatigue and inattention: implications for countermeasures. *International Journal of Vehicle Design, 42*, 67-86.

Thiffault, P., & Bergeron, J. (2003a). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention, 35*, 381-391.

Thiffault, P., & Bergeron, J. (2003b). Fatigue and individual differences in monotonous simulated driving. *Personality and Individual Differences, 34*, 159-176.

Tietze, H., & Hargutt, V. (2001). Zweidimensionale Analyse zur Beurteilung des Verlaufs von Ermüdung. *43. Tagung experimentell arbeitender Psychologen, Regensburg, Germany*. Abgerufen am 19.01.2010 von http://www.psychologie.uni-wuerzburg.de/methoden/texte/2001_tietze_hargutt_Zweidimensionale_Analyse.pdf

Ting, P.-H., Hwang, J.-R., Doong, J.-L., & Jeng, M.-C. (2008). Driver fatigue and highway driving: a simulator study. *Physiology & Behavior, 94*, 448-453.

Tran, Y., Wijesuriya, N., Tarvainen, M., Karjalainen, P., & Craig, A. (2009). The relationship between spectral changes in heart rate variability and fatigue. *Journal of Psychophysiology, 23(3)*, 143-151.

Vandewalle, G., Middleton, B., Rajaratnam, S. M. W., Stone, B. M., Thorleifsdottir, B., Arendt, J., & Dijk, D.-J. (2007). Robust circadian rhythm in heart rate and its variability: influence of exogenous melatonin and photoperiod. *Journal of Sleep Research, 16*, 148-155.

Van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *Journal of Neuroscience, 28(8)*, 1816-1823.

Verwey, W. B., & Zaidel, D. M. (2000). Predicting drowsiness accidents from personal attributes, eye blinks and ongoing driving behaviour. *Personality and Individual Differences, 28*, 123-142.

Wertheim, A. H. (1991). Highway hypnosis: a theoretical analysis. In A. D. Gale, I. Brown, C. M. Haslegrave, I. Moorhead, & S. P. Taylor (Eds.), *Vision in Vehicles III (pp. 467-472)*. North Holland: Elsevier.

Wierwille, W. W., & Ellsworth, L. A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis and Prevention, 26*, 571-581.

Wilhelm, B. J. (2008). Pupillographie zur Messung von Fahrerschläfrigkeit. *Klinische Monatsblätter für Augenheilkunde, 225*, 791-798.

Williams, H. L., Lubin, A., & Goodnow, J. J. (1959). Impaired performance with acute sleep loss. *Psychological Monographs: General and Applied, 73(14)*, 1-26.

Wright, N. A., & McGown, A. S. (2001). Vigilance on the civil flight deck: incidences of sleepiness and sleep during long-haul flights and associated changes in physiological parameters. *Ergonomics, 44*, 82-106.

Wright, N. A., Stone, B. M., Horberry, T. J., & Reed, N. (2007). *A review of in-vehicle sleepiness detection devices. Published Project Report 157*. TRL Limited. Abgerufen am 19.01.2010 von http://www.masstransitmag.com/pdf/2007/mar/ABMT328Sleepy.pdf

Yoss, R. E. (1969). The sleepy driver: a test to measure ability to maintain alertness. *Mayo Clinic Proceedings, 44(11)*, 769-83.

Einzelarbeiten

Studie 1:

Simon, M., **Schmidt, E. A.**, Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., Bogdan, M., Rosenstiel, W., & Schrauf, M., EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Manuscript submitted for publication*.

Studie 2:

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis and Prevention*, 41, 1087-1093.

Studie 3:

Schmidt, E. A., Schrauf, M., Simon, M., Buchner, A., & Kincses, W. E. The short-term effect of verbally assessing drivers' state on vigilance indices during monotonous daytime driving. *Manuscript submitted for publication.*

Auswahl weiterer Publikationen des Autors aus der Promotionszeit:

Schmidt, E.A., Kincses, W.E., Schrauf, M., Haufe, S., Schubert, R., & Curio, G. (2007). Assessing Drivers' Vigilance State During Monotonous Driving. *Proceedings of the Fourth International Symposium on Human Factors in Driving Assessment, Training, and Vehicle Design, Stevenson, WA,* 138-145.

Kincses, W.E., Hahn, S., Schrauf, M., & **Schmidt, E.A.** (2008). Fahrerassistenzsysteme - Mobiles EEG zur Messung der mentalen Fahrerbeanspruchung. *Automobiltechnische Zeitschrift (ATZ), 110(3),* 204-209.

Schmidt, E.A., Tischler, M.A., Schrauf, M., & Kincses, W.E. (2009). Die physiologische Erfassung mentaler und physischer Beanspruchung bei Nachtfahrten mit Lichtassistenzsystemen. *Der Fahrer im 21. Jahrhundert, VDI-Bericht 2085,* 181-192.

EEG Alpha Spindle Measures as Indicators of Driver Fatigue Under Real Traffic Conditions

Corresponding author:

Michael Simon [a,c] *michael.m.simon@daimler.com*

phone: +49 (0) 731 505 2454; fax: +49 (0) 731 505 4105

Eike A. Schmidt [b], Wilhelm E. Kincses [a], Martin Fritzsche [a], Andreas Bruns [a,d], Claus Aufmuth [a], Martin Bogdan [c], Wolfgang Rosenstiel [c], Michael Schrauf [a,b]

[a] Daimler AG; Research and Development; HPC H602; Leibnizstr. 2; D-71032 Böblingen

[b] Heinrich-Heine-University Düsseldorf; Institut für Experimentelle Psychologie; Universitätsstr. 1; D-40225 Düsseldorf

[c] University of Tübingen; Wilhelm-Schickard-Institut für Informatik; Sand 13; D-72076 Tübingen

[d] Current address: F. Hoffmann-La Roche AG, PRDN-T3, Grenzacherstrasse 124, 4070 Basel, Switzerland

Acknowledgments

This research was supported by the German Federal Ministry of Education and Research (BMBF) Grant 16SV2233 awarded to Wilhelm E. Kincses. We thank Sven Willmann for the technical support.

Keywords: EEG; alpha rhythm; alpha spindle, fatigue detection; driver monitoring

Abstract

Objective: The purpose of this study is to show the effectiveness of EEG alpha spindles, defined by short narrowband bursts in the alpha band, as an objective measure for assessing driver fatigue under real driving conditions.

Methods: An algorithm for the identification of alpha spindles is described. The performance of the algorithm is tested based on simulated data. The method is applied to real data recorded under real traffic conditions and compared with the performance of traditional EEG fatigue measures, i.e. alpha band power. As a highly valid fatigue reference, the last 20 min of driving from participants who aborted the drive due to heavy fatigue were used in contrast to the initial 20 minutes of driving.

Results: Statistical analysis revealed significant increases of several alpha spindle parameters and among all traditional EEG frequency bands, only of alpha band power; with larger effect sizes for the alpha spindle based measures. An increased level of fatigue over the same time periods for drop-outs, as compared to participants who did not abort the drive, was observed only by means of alpha spindle parameters.

Conclusion: EEG alpha spindle parameters increase both fatigue detection sensitivity and specificity as compared to EEG alpha band power.

Significance: It is demonstrated that alpha spindles are superior to EEG band power measures for assessing driver fatigue under real traffic conditions.

Introduction

Extensive research has identified fatigue as a major problem in safety critical work situations as well as in public traffic (Dinges, 1995; Folkard, 1997). Despite the methodological difficulties of reliably assessing the causes for fatal traffic accidents, various studies conclude that 15-20 % of all fatal traffic accidents are caused by driver fatigue (Horne and Reyner, 1999; Dawson et al., 2000; Hell and Langwieder, 2001; Connor et al., 2002; NHTSA, 2003). For the European Union this translates into approximately 6,000 fatigue-related traffic fatalities which still occur each year despite the constant decline of overall fatalities on European roads over the past twenty years (CARE, EU road accidents database, 2009).

The automotive industry follows various approaches to increase traffic safety by assessing driver fatigue and by taking corresponding actions, e.g. informing and warning the driver accordingly. Based on the assumption that fatigue is accompanied by behavioral correlates reflected in the driver's vehicle control, currently available in-vehicle fatigue detection systems are based mainly on driving performance data such as steering wheel activity, lane keeping performance, etc. Examples of such systems are "Attention Assist" (Mercedes-Benz) and "Driver Alert Control" (Volvo). In order to develop and validate these rather indirect measures, an objective criterion is required to reliably measure driver fatigue. Following the requirements of test theory and the applicability in the driving context, an ideal fatigue measure should fulfill several conditions. It has to be specifically sensitive to variations of fatigue (*specificity* and *sensitivity*) and has to reliably and consistently assess the same state over various measurements (*reliability*). Especially the requirement of specificity is crucial when conducting in-car experiments under real traffic conditions, since unconstrained experimental conditions as well as a much higher level of noise and artifacts in comparison with laboratory conditions are to be expected. Furthermore, the measure has to provide an adequate *temporal resolution* that accounts for the dynamics of the fatigue process, which was shown to exhibit pronounced fluctuations, especially when participants experience microsleep episodes or fight against sleep (Harrison and Horne, 1996; Lal and Craig, 2005). In order to provide a maximum of ecological validity we conducted measurements under real traffic conditions where *unobtrusiveness* and therefore a minimum of interference with the fatigue state is a further essential condition. Finally, the measure's *objectivity*, i.e. its resistance to intended or unintended manipulation poses an important issue, especially with regard to safety critical situations such as driving.

With respect to its causal factors, driver fatigue can be subdivided into sleep related and task related fatigue (May and Baldwin, 2009). Our aim is to measure any form of driver fatigue, regardless of the ultimate causes, since they all produce similar decrements in driving performance culminating in a safety critical level. Furthermore, the detection process may be very similar for all types of fatigue (May and Baldwin, 2009).

Fatigue measurements mostly rely on participants' self reports which are based on various single- and multi-item questionnaires and which have been developed and correlated with both physiological and performance measures (i.e. Stanford Sleepiness Scale “SSS”, Hoddes et al., 1973; Karolinska Sleepiness Scale “KSS”; Åkerstedt and Gillberg, 1990; Kaida et al., 2006). Although of great value for studying the acceptance of a fatigue detection system, drivers' subjective rating scales exhibit limited objectivity as they are vulnerable to manipulation, to memory effects or to insufficient self introspection skills. Furthermore, subjective measures embrace an irresolvable trade-off between retrieval frequency and interference with the fatigue process under observation. Additionally, several on-road and simulator studies have shown that fatigue self assessment might be inaccurate (Belz et al., 2004; Moller et al., 2006; Schmidt et al., 2009). Alternatively, secondary task performance measures have been proposed to assess fatigue. There is contradictory evidence whether such performance measures accurately render the fatigue process. Some authors argue that instead of continuously following its course, it oftentimes comes to a sudden performance breakdown once a certain level of fatigue is reached (Riemersma et al., 1977; Belz et al., 2004).

In contrast, neurophysiology based measures can provide an objective and direct characterization of the driver's cognitive state with high temporal resolution (Lin et al., 2005; Trejo et al., 2007; Shen et al., 2008). Currently, advances in technology and in signal processing enable real-time measurements of drivers' cognitive states under real traffic conditions (Kohlmorgen et al., 2007; Dixon et al., 2009).

Here we propose an electroencephalography (EEG) based fatigue measure that largely fulfills the requirements formulated above and outperforms traditional EEG-based fatigue measures with respect to sensitivity and noise susceptibility.

Electroencephalography (EEG)

In studies inside and outside the driving context, various frequency band power measures (delta, theta, alpha, beta) and combinations such as (theta + alpha)/beta showed fatigue related effects (i.e. Lal and Craig, 2002; Ryu et al., 2007; for reviews see Lal and Craig, 2001 and Oken et al., 2006). As one of the most prominent EEG-indicators thereof, alpha band power (7 - 13 Hz) has been shown to detect early stages of fatigue and was applied in a variety of real-traffic and driving simulator studies (O'Hanlon and Kelly, 1977; Lemke, 1982; Brookhuis and De Waard, 1993; Kecklund and Åkerstedt, 1993; Lal and Craig, 2002; Campagne et al., 2004; Horne and Baulk, 2004; Otmani et al., 2005).

In addition to its significance in fatigue assessment, alpha activity is also suggested to represent a cortical idling rhythm involved in the perceptual process (Pfurtscheller et al., 1996; Klimesch et al., 2007). For example, it could be shown that the detection performance and reaction time for visual stimuli deteriorates with an increase in parieto-occipital alpha activity (Hanslmayr et al., 2007; van Dijk et al., 2008). As driving is mainly a visual task (Lansdown, 2001) the alpha band seems best suited to assess changes in visual perception performance, regardless of their cause.

EEG recordings from drivers in real traffic conditions are much more influenced by artifacts and noise than recordings performed in the laboratory. Artifacts may originate from biological and technical noise inherent to these measurement conditions and disperse their energy over a wide range of frequencies (Goncharova et al., 2003). Consequently, changes in alpha band power alone can not be unambiguously attributed to neural factors, an instance which calls for a more specific approach to distinguish between neural and exogenous causes. Our approach to tackle this problem consists in focusing on spectral microstructures within the alpha band which have been shown in various studies to be indicative for attention and fatigue (Kecklund and Åkerstedt, 1993; Petsche et al., 1997; Cantero and Atienza, 2000; Cantero et al., 2002).

Alpha Spindles

With advancing fatigue, alpha power density increases by means of short (500 ms to several seconds) narrowband "bursts" of alpha band activity (Caille and Bassano, 1977; Kecklund and Åkerstedt, 1993; Tietze and Hargutt, 2001; Cantero et al., 2002; Eoh et al., 2005; Papadelis et al., 2007), which we call *alpha spindles* (Figure 1). These spectral microstructures can be characterized by a narrow frequency peak within the alpha band (7 - 13 Hz) and a low frequency modulation envelope, which results in the typical "waxing and waning" of the alpha rhythm (Shaw, 2003). We hypothesize that alpha spindles represent a characteristic, subject specific fatigue measure described by spectral amplitude, peak frequency and duration which are, due to their narrowband spectral character, less noise susceptible than other, e.g. band power related driver fatigue measures.

The current work focuses on two main aspects. Firstly, we will describe an approach to reliably detect alpha spindles and to describe their characteristics (peak frequency, amplitude, duration) and secondly, using data from real traffic experiments we will demonstrate the advantages of the proposed method as opposed to the established alpha power measure with regard to real traffic applicability and fatigue sensitivity. In a first step we limit our comparison to alpha power because it is the most widely used EEG fatigue measure in the driving context.

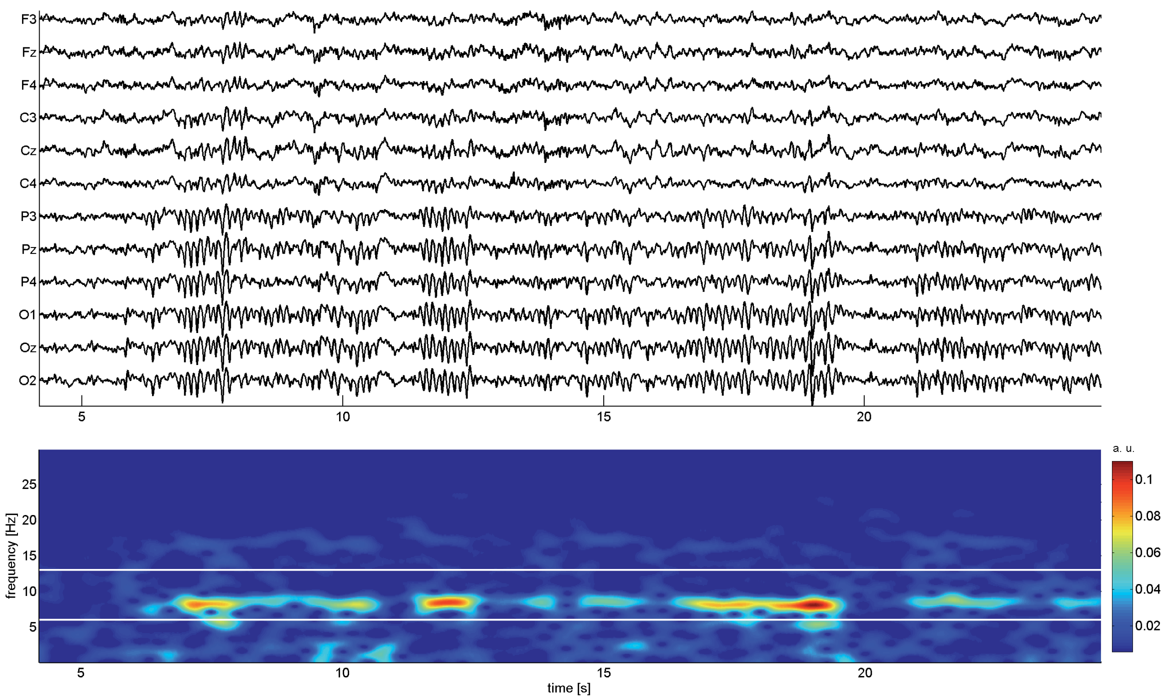

Figure 1: Example of alpha-spindle activity. Spindles are dominant at parieto-occipital electrodes with decaying activity towards frontal electrodes. Clearly visible is the burst-like structure and the low-frequency modulation envelope of the alpha spindles. The bottom plot shows the spectrogram of channel Oz with the alpha band boundaries (7–13 Hz) marked by the white lines. Each spindle has a constant narrow peak frequency range, which can vary between spindles.

Methods

Alpha Spindle Detection

To account for non-stationary characteristic of the EEG as well as to facilitate a computationally efficient implementation for real-time applications, we use the short time Fourier transform (STFT, Oppenheim and Schafer, 1989) in its FFT implementation (Fast Fourier Transformation) for the spectral decomposition of the EEG signal.

Each EEG-channel is divided into segments of 1 s length with an overlap of 750 ms. Subsequently, each segment is made zero mean and multiplied with a Hamming window, the FFT is computed and the spectral amplitude density maximum between 3 and 40 Hz is identified. If this maximum lies within the alpha band (7–13 Hz), the *full width at half maximum* (FWHM) of the spectral peak is determined. If the FWHM is smaller than twice the bandwidth of the Hamming window (indicating a component of sufficiently narrow bandwidth to be considered oscillatory), the time segment undergoes further analysis.

To account for the 1/f-like noise of EEG recordings (Pereda et al., 1998; Wagenmakers et al., 2004), for each EEG channel an exponential curve is fitted to the mean amplitude spectrum obtained as the average across all single-segment amplitude spectra from the entire recording (Figure 2), allowing to separate between "signal" (area above fitted curve) and "noise" (area below fitted curve). Adaptation to varying noise levels over time is achieved by multiplying the exponential fit with the ratio between the integrated spectrum of the current segment and the integrated mean spectrum.

We consider alpha spindles to reflect periods of amplitude modulated cortical steady state activity which (a) can be detected above a certain signal-to-noise ratio (SNR), (b) can be restricted to single electrodes, (c) have a stable peak frequency and (d) can persist for up to several seconds.

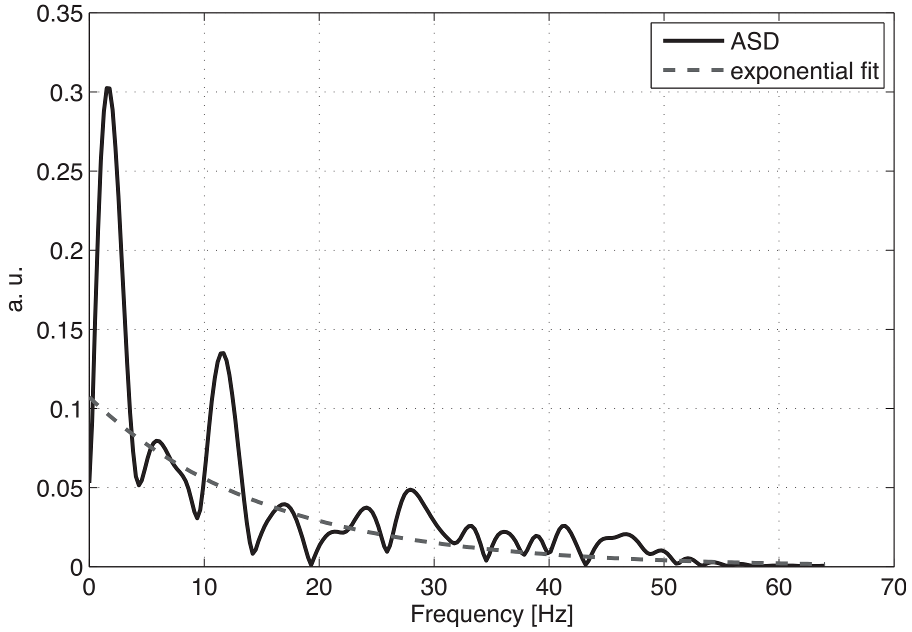

Figure 2: Amplitude spectral density of a 1 s EEG segment with an alpha spindle spectral peak at about 11.5 Hz and the exponential fit as a dashed line. The oscillation index is defined as the ratio of the total area under the peak and the area under the exponential curve.

Consequently, for each EEG channel and for each one-second time epoch, we test whether the total area under the peak, bounded by the FWHM, is at least twice as large as the area in the same frequency interval below the fitted noise line (effectively establishing a signal-to-noise-ratio threshold). We call this ratio the *oscillation index*. Consecutive time segments fulfilling this criterion and having a peak frequency within a 10 % change to the previous segment are grouped to one alpha spindle. The *frequency* of an alpha spindle is calculated as the mean of the peak frequency of all contributing segments; the same applies to the *spectral amplitude*. The *duration* of a spindle is defined as the time span from the start of the spindle's first time segment to the end of its last segment.

Since alpha spindles are discrete events characterized by their *duration, spectral amplitude* and *peak frequency*, we also use the moving average of these parameters over larger time windows to obtain continuous measures better suited to reproduce continuous fatigue changes. Additionally this enables us to count the occurrence rate of alpha spindles within the moving window. We call this parameter the *alpha spindle rate*.

Data

Synthetic Data

To objectively assess the accuracy and the noise susceptibility of the proposed alpha spindle detection, we generated synthetic EEG data with defined spindle SNRs at a sampling rate of 128 Hz. For the background noise we generated a random time-domain signal with a 1/f-like noise spectrum (Wagenmakers et al., 2004) and then added randomly scattered, non-overlapping segments of sinusoids with uniformly distributed random frequencies (7.5–12.5 Hz), phases (0–2π) and durations (.5–3.5 s). All parameters were varied independently of each other. The amplitude of each spindle was adjusted so that it had a predefined SNR relative to the local noise level. At the start and end of a spindle a cosine taper of 100 ms duration was multiplied to obtain a smooth decay towards the boundaries of the spindle.

We simulated 50 channels independently with 400 spindles each, distributed in about 60 min of synthetic EEG data per channel. Simulations were performed independently for SNRs of -6 dB, -3 dB and 0 dB.

Real Data

To test the applicability of the proposed method under real traffic conditions and in order to compare its fatigue sensitivity performance with traditional alpha band power measures, we applied both methods to EEG data sets recorded during two prolonged monotonous daytime driving studies (Schmidt et al., 2009; Schmidt et al., in preparation). These studies proved to successfully have induced driver fatigue and in a subset of participants even severe drowsiness. The experiments took place in a Mercedes-Benz S-Class (W221). To create a monotonous driving situation, a low traffic highway (A81) in Germany was chosen. If completed, the length of the drive was about 430 km for the first study and about 480 km for the second study. While driving, the participants completed a monotonous secondary reaction time task (auditory odd-ball paradigm) which is not analyzed in this study (for details see Schmidt et al. 2009). All driving sessions started at about 12:45 pm. For safety reasons an investigator accompanied the drive at all times. Participants were instructed not to exceed a speed of 130 km/h, not to talk to the investigator and not to use any in-car devices (radio, cruise control, etc.). The participants were free to abort driving at any time without any disadvantages. For a detailed description of the experiment please refer to the initial work of Schmidt et al. (2009).

Altogether, 10 out of 55 participants (6 male, 4 female; age: M = 27.5; range: 24-36) aborted their drive due to severe fatigue (average driving time: 2:23 h, standard deviation (SD): 0:38 h, range: 0:56 – 3:15 h), which provides the most objective fatigue criterion available. Hence for the following analyses (unless stated differently) we only used the data from these ten participants (called "drop-outs" in the following). Especially for these subjects, the contrast in driver fatigue between sections at the beginning and at the end of the drives should be maximal since the study was performed at daytime with rested participants at the start of their drive. This is also underlined by the results of the drop-outs' self rating on the Karolinska Sleepiness Scale (KSS; Åkerstedt and Gillberg, 1990), ranging from one (extremely alert) to nine (extremely sleepy, fighting sleep) that was assessed every 20 minutes throughout the drive. The average KSS value for the first 20 minutes was 4.3 (SD = 1.9) and increased to 8.5 (SD = .5) before break-off. We therefore used the first and the last 20 minute period of each drive here.

EEG was recorded with BrainAmp hard- and software (Brain Products GmbH, Munich, Germany) with 128 electrodes and 1000 Hz sampling rate for study one and 64 electrodes and 500 Hz sampling rate for study two. Electrode placement was based on the extended international 10-20 system. Both studies used the nose bridge as reference. For further analysis we used the overlapping set of 64 electrodes from both studies. Data were band-pass filtered from .5 Hz to 48 Hz and down-sampled to 128 Hz. In order to minimize the influence of muscle activity, eye blinks and technical noise, we used the extended infomax ICA algorithm (Lee et al., 1999), available in the EEGLab toolbox (Delorme and Makeig, 2004), to reject artifactual components. Data from 43 out of 62 EEG channels were used; temporal channels were excluded from analyses due to prolonged interference with muscle activity; channels Cz and FC2 had to be excluded due to technical problems with the EEG hardware.

The proposed algorithm for alpha spindle detection was used to identify alpha spindles and to calculate spindle rate, average spindle amplitude, duration and frequency for both 20 minute driving sections. Alpha band power (7–13 Hz) was calculated using Welch's modified periodogram.

Statistical Analysis

Synthetic Data

Synthetic data was analyzed by means of ROC (receiver operating characteristic) curves at different SNRs. The signal was segmented in overlapping time segments in the same manner as described for the alpha spindle detection (1 s segments, 750 ms overlap). A segment that overlapped to at least 50 % with a simulated spindle was assigned to the positive class. Otherwise it was assigned to the negative class. A segment was counted as "true positive" if a spindle detected by the above described algorithm overlapped by at least 50% with a positive class segment. A "false positive" was counted if a detected spindle overlapped by at least 50% with a negative class segment. The true positive rate was defined as the sum of true positives divided by the total number of positive class segments. The false positive rate was defined as the sum of false positives divided by the total number of negative class segments. The threshold systematically varied along the ROC curves was the oscillation index.

Real Data

Our main focus was on comparing the EEG parameters from the first and last 20 min driving sections, respectively. For each subject, spindle rate, mean spindle amplitude, duration and frequency as well as mean alpha-band power were obtained for each of the 20 min time segments and then further averaged within each of three channel groups: frontal (F: 3, 1, z, 2, 4; FC: 3, 1, z, 4), central (C: 3, 1, 2, 4; CP: 3, 1, z, 2, 4) and parieto-occipital (P: 3, 1, z, 2, 4; PO: 3, z, 4; O: 1, z, 2). Thus, each parameter yielded 6 values for each subject (2 driving sections · 3 channel groups), which were analyzed using a 2-way repeated-measures design (separately for each parameter) with within-subject factors "driving section" and "channel group". The repeated-measures design was implemented via a multivariate (MANOVA) approach in order to circumvent problems arising from violations of the sphericity assumption for the factor "channel group" (O'Brien and Kaiser, 1985). The test criterion reported is the (exact) F-statistic derived from Pillai's trace; the level of α was set to .05 for all analyses. The partial η^2 is reported as a measure of relative effect size whenever H_0 had to be rejected. For statistically significant results of the main effect "channel group" a post-hoc analysis was applied in which each factor level was compared to the previous level.

Results

Synthetic Data

In order to establish a relationship between the SNR of a signal and the oscillation index of our proposed alpha spindle detection algorithm, we calculated the average oscillation index for all correctly detected spindles in the simulated data. The results are presented in Table 1. The average oscillation index for the real EEG data from the participants who aborted the drive was 3.9 (SD = .6). Therefore a SNR of -3 dB in our simulation corresponds approximately to the same noise level observed in the real data.

For a SNR of -3 dB, we can achieve a sensitivity of more than 95 % while having less than 3 % false-positive rate.

Figure 3 shows the performance of the alpha spindle detection for different signal-to-noise ratios as ROC-curves. As expected, the performance increases with the SNR with converging improvement for increasing SNR.

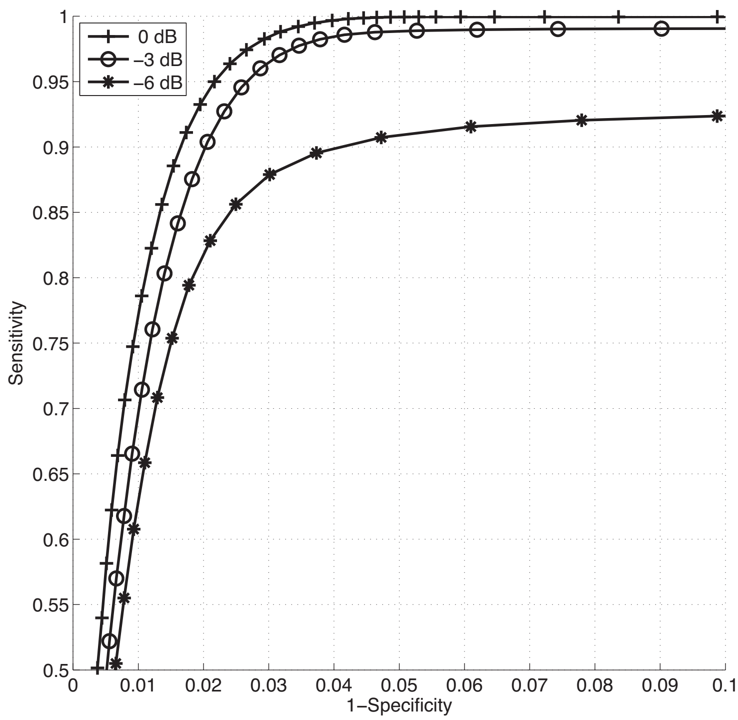

Figure 3: ROC curves for the alpha spindle detection performance for simulated data. Curves for –6 dB, -3 dB and 0 dB SNRs are shown.

Table 1 further compares the accuracy of alpha spindle parameter estimation for all SNRs by computing the Root Mean Squared Error (RMSE) for all correctly detected spindles. As the algorithm's frequency detection error (cf. Table 1) is within the frequency resolution of the FFT, we can precisely locate the position of the synthetic alpha peaks in the frequency domain. The error in the estimation of the spindle duration is mainly due to the decay of the simulated spindles to zero at the boundaries, which makes it difficult to assess the exact duration. The error of the spectral amplitude is strongly influenced by the bias and variance of the spectral estimation methods (Oppenheim and Schafer, 1989), but still sufficiently low. The accuracy of parameter estimation improves with increasing signal strength, but is sufficient for all SNRs.

Table 1: Accuracy of parameter estimation for alpha spindles using the algorithm's default settings described in the text. Mean spindle duration was 2 s and mean spindle frequency 10 Hz. Since the spindle amplitude was dependent on the respective SNR, the actual mean value is given in brackets. Presented are the combined results for 50 channels with 400 simulated spindles each. RMSE stands for Root Mean Squared Error.

	SNR [dB]		
	-6	-3	0
Oscillation index	3	3.8	4.6
RMSE frequency [Hz]	.06	.05	.04
RMSE duration [s]	.21	.19	.18
RMSE amplitude [a.u.]	0.003 (0.024)	0.0028 (0.033)	0.003 (0.048)

Real Data

The results of the repeated-measures analysis of variance on the effects of driving section and channel group are summarized in Table 2. Figure 4 visualizes the results by showing the means over all participants for each dependent measure, channel group and driving section.

Table 2: Statistical results for the repeated-measures analysis (MANOVA) of driving section and channel group for alpha spindle measures and alpha band power.

	Main Effect						Interaction		
	driving section			channel group			driving section x channel group		
	$F(1,9)$	p	η2	$F(2,8)$	p	η2	$F(2,8)$	p	η2
Spindle Rate	22.56	.001	.715	.803	.481	n.s.	.54	.601	n.s.
Spindle Duration	21.28	.001	.703	1.05	.393	n.s.	.65	.550	n.s.
Spindle Amplitude	9.67	.013	.518	1.57	.266	n.s.	.58	.581	n.s.
Spindle Frequency	.04	.849	n.s.	16.20	.002	.802	.60	.572	n.s.
Alpha Power	8.18	.019	.476	1.41	.299	n.s.	.48	.638	n.s.

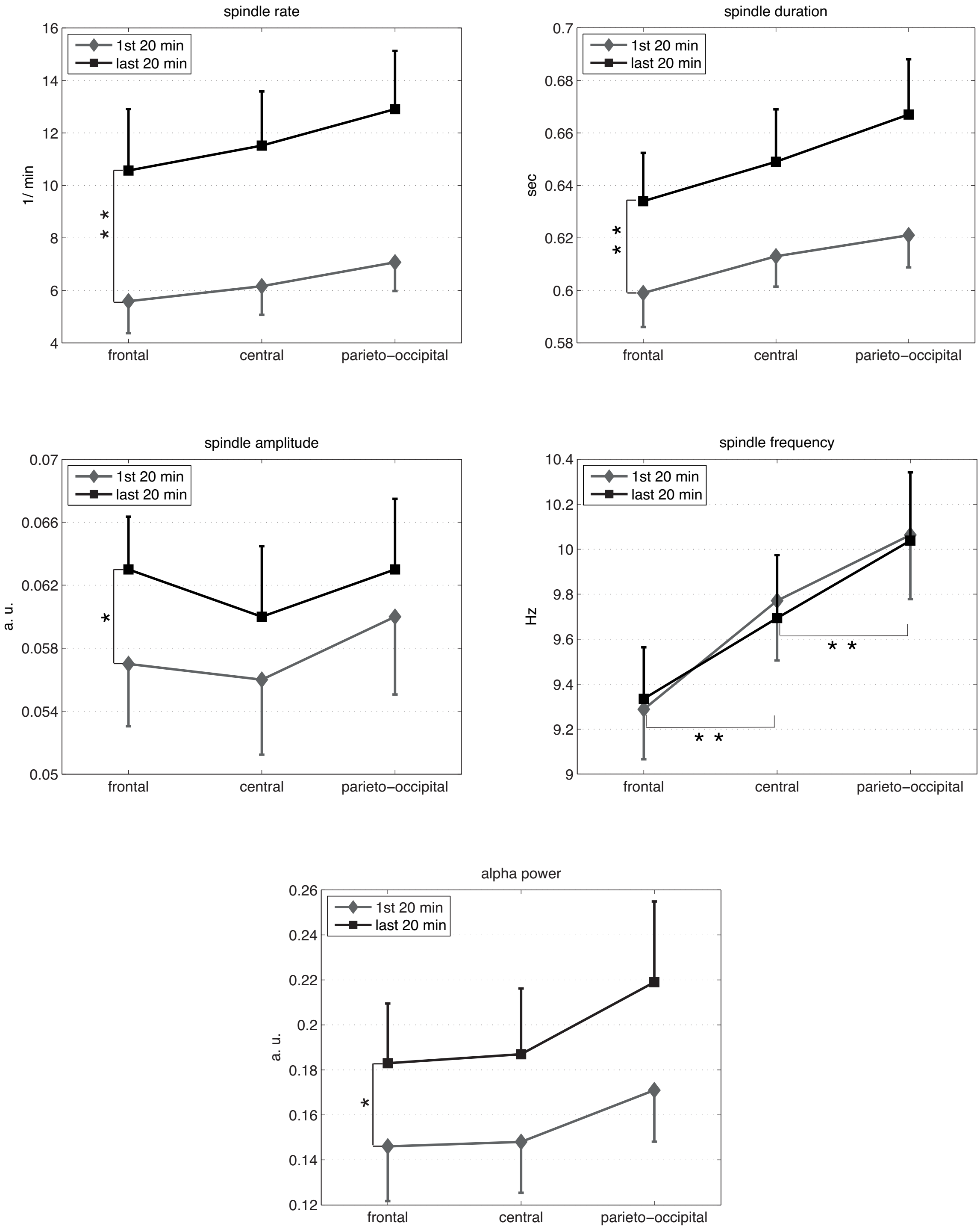

Figure 4: Comparison of the first and last section of the drive for three channel groups. Values represent averages over all participants. Error bars indicate the standard errors of the means (* $p < .05$, ** $p < .01$).

All measures except spindle frequency showed significant differences ($p < .05$) between the first and last driving epoch. Spindle rate, spindle duration and spindle amplitude showed the highest differences between the two driving sections as reflected in the partial η^2-values, with the highest η^2-value (.715) for spindle rate.

Only spindle frequency showed a significant effect for the channel groups. Frontal channels had a lower frequency than central channels ($F(1,9) = 20.55$; $p = .001$; $\eta^2 = .695$) and central channels were lower in frequency than the parieto-occipital group ($F(1,9) = 19.62$; $p = .002$; $\eta^2 = .686$). Figures 4a-d indicate a tendency of increased alpha spindle activity going from anterior to posterior sites, reflected by a higher spindle rate and longer spindle duration.

For none of the dependent measures a significant interaction of driving section and channel group was observed, indicating similar changes in parameter values with increasing fatigue for all channel groups.

Considering the relative increase of the fatigue measures from the first to the last driving section, spindle rate (Figure 4a) showed the highest dynamic with 90% increase for all channels as compared to only 32% increase for alpha power (Figure. 4e).

Drop-Outs / Non Drop-Outs Comparison

In the following “drop-outs” refers to participants who aborted the drive due to excessive fatigue and “non drop-outs” to those who completed the whole driving course.

For all participants the devolvement of fatigue was confounded with the factor time-on-task. To disambiguate the influence of the two factors fatigue and time, we compared the data of non drop-outs and drop-outs from the same experiments. 31 valid non drop-out data sets were available. For those we limited the driving duration to a randomly assigned driving duration sampled from a normal distribution with mean 2:23h and standard deviation 0:38 h (same distribution of driving duration as for the drop-outs).

We only considered parameters that had significant effects in the preceding analysis (all measures except the spindle frequency). We entered the mean of each parameter in the first and last 20 min of driving averaged over all channels into a 2-way repeated-measures analysis for within-subject (“driving section”) and between-subjects (“drop-out”) comparisons (again implemented using a MANOVA approach). Averaging over channels was reasonable since no significant interaction between driving segment and electrode group was found. The results are presented in Table 3 and Figure 5.

Table 3: Statistical results for the comparison of drop-outs vs. non drop-outs using a repeated-measures analysis with within-subject factor "driving section" and between-subjects factor "drop-out". A multivariate approach (MANOVA) was used.

	Main Effect						Interaction		
	driving section			drop-out			driving section x drop-out		
	$F(1,39)$	p	η^2	$F(1,39)$	p	η^2	$F(1,39)$	p	η^2
Spindle Rate	58.95	<.001	.602	11.56	.002	.229	8.30	.006	.175
Spindle Amplitude	17.96	<.001	.315	.01	.923	n.s	8.54	.006	.180
Spindle Duration	34.45	<.001	.469	11.33	.002	.225	6.65	.014	.146
Alpha power	11.52	.002	.228	.141	.710	n.s.	3.52	.068	n.s.

We observed significant within-subject effects of driving section for all dependent measures reflecting an increase from the first to the last 20 minutes of driving. Further, there were significant interactions between driving section and subject group for the dependent measures spindle rate, spindle amplitude and spindle duration indicating a larger increase of the dependent measures from early to late sections for the drop-outs as compared to the non-drop-outs. There were also significant between-subject effects for spindle rate and spindle duration with higher levels for drop-outs.

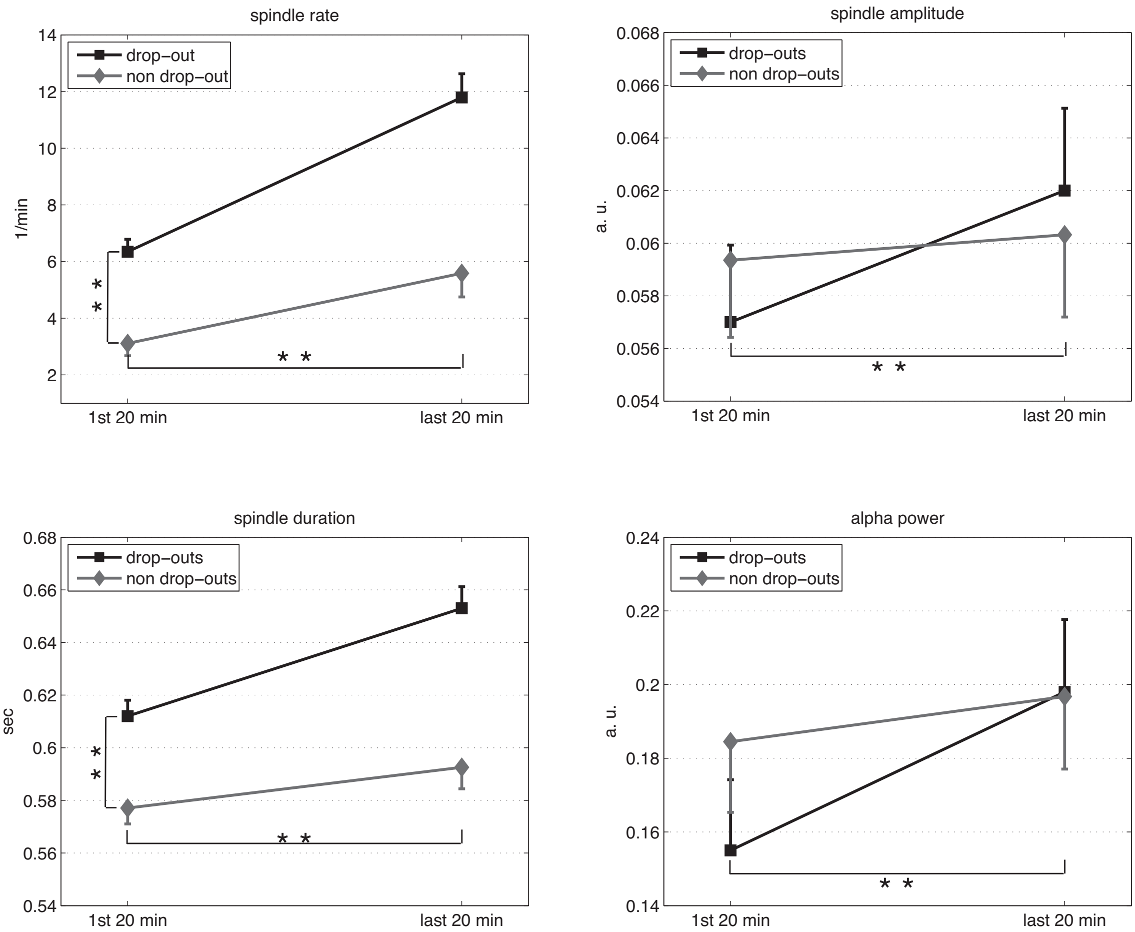

Figure 5: Mean parameter values for the comparison of drop-outs and non drop-outs. Error bars indicate the standard errors of the means (* $p < .05$, ** $p < .01$).

Temporal Resolution

So far, only mean values of the 20 min sections from the beginning and the end of a drive have been reported. To address a higher temporal resolution of the fatigue measures, mean values of windows of 1 to 10 min length within the 20 min sections were computed and the discriminative power was analyzed. Only alpha spindle rate was considered, since this has proven to be the most discriminative parameter in the previous analysis. A ROC analysis of the first (awake) vs. the last 20 minutes (fatigued) of driving was conducted. This ROC analysis is different to the one conducted for the simulated spindles. It investigates how accurately an alpha spindle rate value, measured on different time scales, can predict the fatigue state of a driver. The true positive rate is the number of true positives divided by the total number of windows from the last 20 min. The false positive rate is the number of false positives divided by the total number of windows from the first 20 min.

The alpha spindle rate was computed for every channel in overlapping time windows of 1 to 10 min duration with a 10 s step width for the first and last 20 min of driving. In order to pool the spindle rate of all drop-outs, the data had to be normalized, since the participants had different offsets and variances in their respective spindle rates. Normalization is done by firstly subtracting the mean of a time window computed over all channels from each channel and secondly, by dividing each channel by the standard deviation computed over all channels. This way the ratios between the channels were preserved. Since no significant differences between the channel groups were observed for spindle rate (see Table 2), we averaged the value of each time window over all channels. Next the data of all drop-outs were pooled.

ROC curves were generated by varying a threshold for spindle rate deciding whether a time window was considered awake or fatigued.

Figure 6 shows the ROC curves for window lengths of 1, 3, 5, 7 and 10 min and the corresponding area under curve (AUC, Fawcett, 2006). All temporal resolutions show a clear ability to distinguish between the two fatigue states, but with varying quality. As one would expect, performance improves with increasing window length at the expense of temporal resolution. Especially from 1 min to 3 min there is a clear performance improvement.

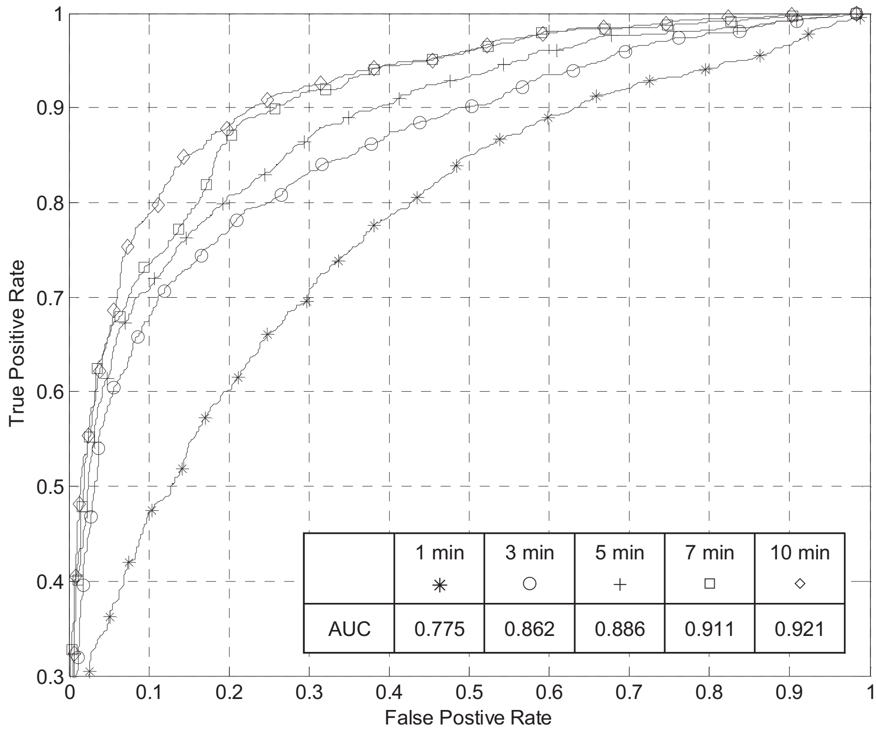

Figure 6: ROC analysis of the spindle rate computed in different window lengths for the first and last 20 min of driving. Discrimination between the two sections was performed. The table shows the area under curve (AUC) for each window length.

Discussion

Based on simulated data we were able to show that the FFT-based detection delivers very good results with regard to detection performance and accuracy of alpha spindle parameter estimation. By approximating the noise level with an exponential fit, we estimated the oscillation index in our experimental data to an average value of 3.9 (SD = .6) which corresponds to a SNR of -3 dB. We further showed on the basis of real EEG data from a real traffic study that the proposed method is better suited to measure driver fatigue under real driving conditions than traditional EEG band power measures.

Statistical analysis of real driving data revealed significant increases between the first (awake) and the last (drowsy) 20 minutes of the drive for all alpha spindle parameters except spindle frequency. For the same alpha spindle parameters effect sizes were higher than for alpha band power (see Table 2). This confirms our hypothesis that alpha spindles are a more sensitive indicator of driver fatigue for real-traffic experiments than alpha power. On group level, we found a significant increase of fatigue-related parameters in general, without being prominent at a particular site. This is in contradiction with studies which showed prominent increase of alpha power over central and parietal sites, but not frontal sites (for reviews see Lal and Craig, 2001 and Oken et al., 2006). However, one possible explanation is the different definition of frequency band boundaries. For example, fatigue related effects at frontal sites were reported for the theta band with an upper boundary of 8 Hz (Lal and Craig, 2002; Strijkstra et al., 2003), which overlaps by 1 Hz with our alpha band starting at 7 Hz. Another reason could be the individual site with prominent alpha spindle activity ranging from parieto-occipital to frontal-midline (see Figure 7). Cortical sites with significant differences that vary individually may result in reduced effects at all sites on group level. One decisive difference between the above mentioned articles and our studies is represented by the fact that our measurements were performed under real driving and real traffic conditions as compared to measurements performed in driving simulators and often under sleep deprivation.

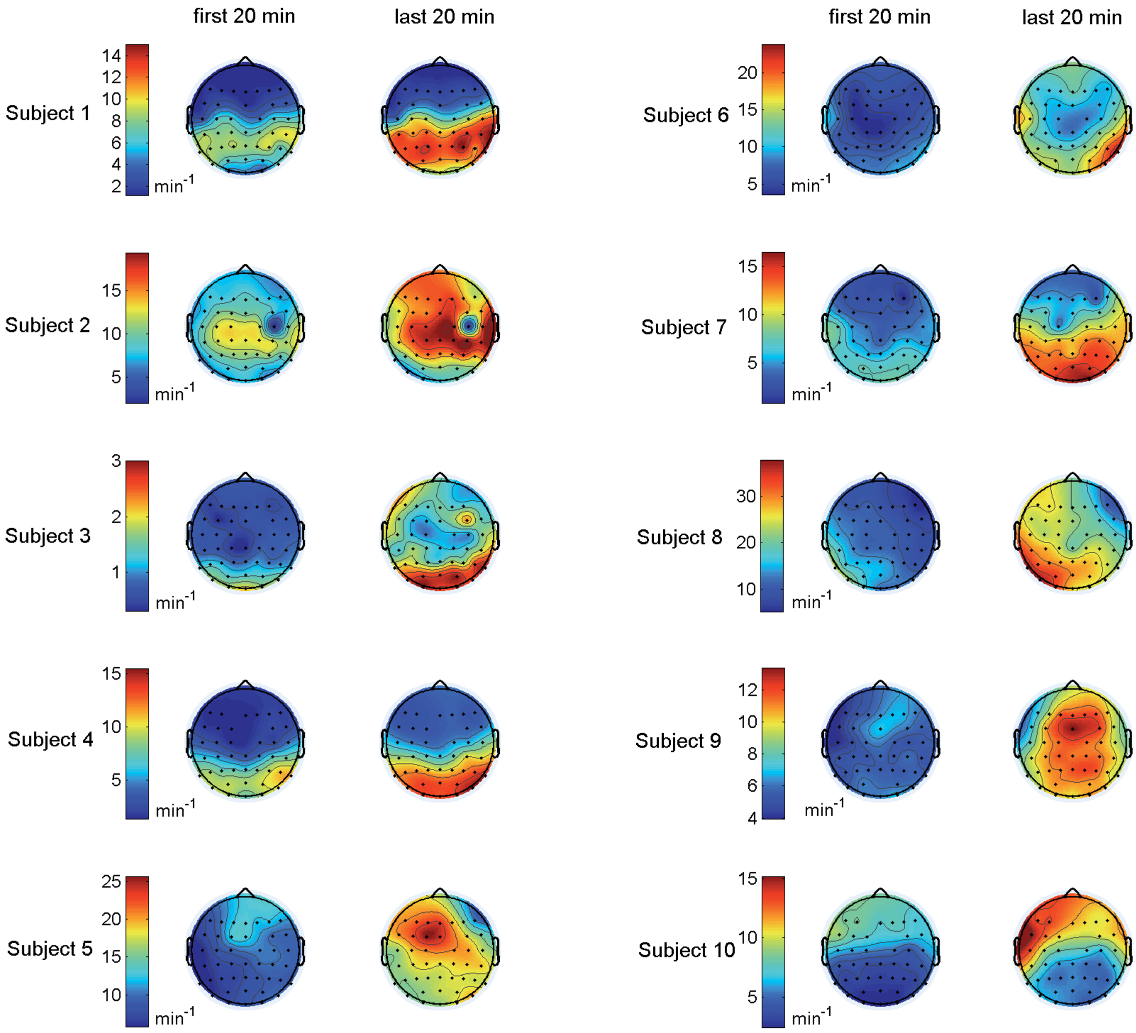

Figure 7: Alpha spindle rate for all drop-out participants from the first and last 20 min of driving. The cortical site with prominent alpha spindle activity varies between subjects.

The comparison of drop-outs and non drop-outs confirmed the adequacy of the alpha spindle parameters as measures for critical fatigue. Since a significant effect of the factor "driving section" (time-on-task) had to be expected as it is confounded with emerging fatigue, the interaction between driving section and subject group revealed a more pronounced parameter increase for spindle rate and spindle duration for participants who aborted the drive (see Figure 5). Consequently, both parameters allow discriminating between mere time-on-task effects and critical fatigue.

Additionally, testing for the simple main effect (two-sample t-test) of the between-subjects factor within the first 20 minutes revealed that spindle rate ($T(39) = 3.41$; $p = .002$) as well as spindle duration ($T(39) = 2.93$, $p = .006$) allow to distinguish between drop-outs and non drop-outs already within the first 20 minutes, which is not possible based on the alpha power measure. A similar study was performed by Otmani et al. (2005), where sleep-deprived and non sleep-deprived participants could not be discriminated at the start of a drive based on theta and alpha band power. This further demonstrates the higher sensitivity of the alpha spindle measures to critical fatigue.

The ROC-analysis of the alpha spindle rate's temporal resolution revealed a trade-off between performance and resolution (Figure 6). The performance improvement for longer time windows is a result of a smaller variance of the data as compared to short time windows. This is caused by two factors. Firstly, longer moving average windows smooth the signal and result in a reduced variance. Secondly, short-time fatigue variations detectable with an accordingly high temporal resolution are not taken into account by the two 20 min sections "awake " and "fatigued" which postulate a constant state over the entire duration. We conclude that for our application domain, a window length of three to five minutes results in an acceptable trade-off between performance and temporal resolution.

In the present study we only aimed at separating two clearly defined distinct states: awake vs. fatigued. Figure 8 demonstrates the alpha spindle rate's ability to reproduce the devolution between these two states in comparison with alpha power. It shows the average of both parameters for the drop-outs in every driving duration quintile of each individual drive. Whereas the spindle rate shows a strictly monotonic, almost linear increase with low standard deviations, the alpha band power shows a much lower increase over time and a reduced consistency across subjects represented in higher standard deviations. This supports our hypothesis that alpha spindles are fatigue indicators with both higher specificity and higher sensitivity than traditional alpha band power measures.

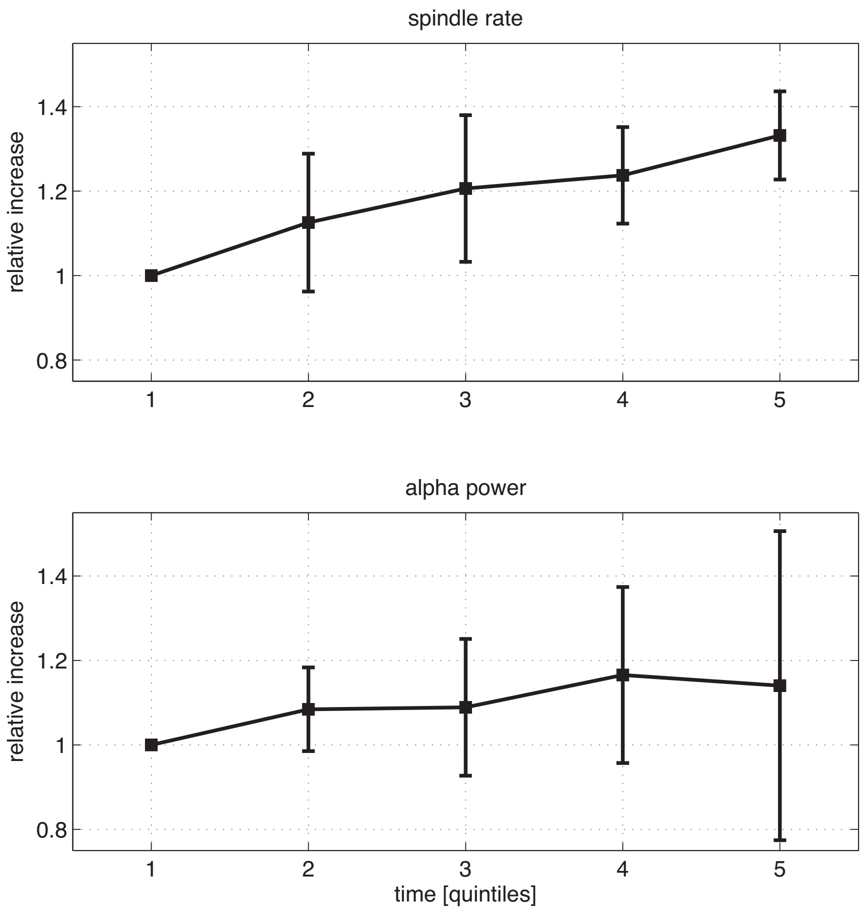

Figure 8: Time evolution of spindle rate and absolute alpha band power. Due to the unequal driving durations, driving time was divided into 5 segments (quintiles) and segments 2 to 5 were referenced to the first segment. The average over all channels was used to calculate the mean and standard deviation over all subjects indicated by error bars.

In our analysis we limited the comparison of alpha spindle parameters to alpha power only because it is the most widely used EEG fatigue indicator in the driving context. For the sake of completeness we also report the results for delta, theta and beta power and also the frequently used ratio (alpha + theta)/beta (power). The same analysis approach as described in the chapter Statistical Analysis was used. Results are shown in Table 4. None of these band power measures shows a significant effect of driving section. This confirms previous findings according to which alpha band based measures are better suited for real traffic experiments than other frequency bands (Caille and Bassano, 1977; O'Hanlon and Kelly, 1977; Kecklund and Åkerstedt, 1993; Lin et al., 2005).

Table 4: Statistical results for EEG band power measures. The same repeated-measures design as described in chapter Statistical Analysis was used.

	Main Effect						Interaction		
	driving section			channel group			Driving section x channel group		
	F(1,9)	*p*	η2	*F*(2,8)	*p*	η2	*F*(2,8)	*p*	η2
Theta Power	.790	.397	n.s.	1.842	.220	n.s.	2.064	.189	n.s.
Beta Power	3.534	.093	n.s.	2.334	.159	n.s.	1.937	.206	n.s.
Delta Power	.575	.468	n.s.	3.287	.091	n.s.	.222	.806	n.s.
(Alpha+Theta)/Beta	2.528	.146	n.s.	.698	.525	n.s.	4.547	.048	.532

References

Åkerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 1990; 52: 29-37.

Belz SM, Robinson GS, Casali, JG. Temporal separation and self rating of alertness as indicators of driver fatigue in commercial motor vehicle operators. Human Factors, 2004; 46: 1, 154-169.

Brookhuis KA, de Waard D. The use of psychophysiology to assess driver status. Ergonomics, 1993; 36: 1099-110.

Caille EJ, Bassano JL. Validation of a behavior analysis methodology: variation of vigilance in night driving as a function of the rate of carboxyhemoglobin. In: Mackie RR, editor. Vigilance. New York: Plenum Press, 1977: 59-72.

Campagne A, Pebayle T, Muzet A. Correlation between driving errors and vigilance level: influence of the driver's age. Physiology & Behavior, 2004; 80: 515-524.

Cantero JL, Atienza M. Spectral and Topographic Microstructure of Brain Alpha Activity During Drowsiness at Sleep Onset and REM Sleep. Journal of Psychophysiology, 2000; 13(3): 151-158.

Cantero JL, Atienza M, Salas RM. Human alpha oscillations in wakefulness, drowsiness period, and REM sleep: different electroencephalographic phenomena within the alpha band. Neurophysiologie Clinique/Clinical Neurophysiology, 2002; 32(1): 54-71.

CARE, EU road accidents database, 2009. Road safety evolution in EU. http://ec.europa.eu/transport/road_safety/observatory/doc/historical_evol.pdf

Connor J, Norton R, Ameratunga S, Robinson E, Civil I, Dunn R, Bailey J, Jackson R. Driver sleepiness and risk of serious injury to car occupants: population based case control study. British Medical Journal, 2002; 324: 1125.

Dawson D, Fletcher A, Hussey F. Beyond the midnight oil. Managing fatigue in transport. Standing Committee on Communications, Transport and the Arts, Australia, 2000.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 2004; 134: 9-21.

Dinges DF. An overview of sleepiness and accidents. Journal of Sleep Research, 1995; 4: 4-14.

Dixon KR, Hagemann K, Basilico J, Forsythe C, Rothe S, Schrauf M, Kincses WE. Improved Team Performance Using EEG- and Context-Based Cognitive-State Classifications for a Vehicle Crew. In: Schmorrow DD et al. (Eds.): Foundations of Augmented Cognition. HCI International 2009 San Diego, USA.

Eoh HJ, Chung MK, Kim SH. Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. Int. Journal of Industrial Ergonomics, 2005; 35: 307-320.

Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters, 2006; 27(8): 861-874.

Fleiss JL. Statistical Methods for Rates and Proportions (2nd ed.). New York: John Wiley & Sons, 1981.

Folkard S. Black times: temporal determinants of transport safety. Accident Analysis and Prevention, 1997; 29(4): 417-430.

Freeman WJ. Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clin Neurophysiol, 2004; 115(9): 2077-2088.

Goncharova II, McFarland DJ, Vaughan TM, Wolpaw JR. EMG contamination of EEG: spectral and topographical characteristics. Clin Neurophysiol, 2003; 114: 1580-1593.

Hanslmayr S, Aslan A, Staudigl T, Klimesch W, Herrmann CS, Bäuml KH. Prestimulus oscillations predict visual perception performance between and within subjects. Neuroimage, 2007; 37: 1465-1473.

Harris FJ. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. Proc. of the IEEE, 1978; 66: 66-67.

Harrison Y, Horne, JA. Occurance of 'microsleeps' during daytime sleep onset in normal subjects. Electroencephalographgy and Clinical Neurophysiology, 1996; 98: 411-416.

Hell, W., Langwieder K. Einschlafunfälle im Straßenverkehr – Eine bisher oft verkannte Unfallursache. Gesamtverband der Deutschen Versicherungswirtschaft, 2001.

Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. Psychophysiology, 1973; 10: 431-436.

Horne J, Reyner L. Vehicle accidents related to sleep: a review. Occupational and Environmental Medicine, 1999; 56: 289-294.

Horne JA, Baulk SD. Awareness of sleepiness when driving. Psychophysiology, 2004; 4: 97-110.

Huupponen E, Gómez-Herrero G, Saastamoinen A, Värri A, Hasan J, Himanen SL. Development and comparison of four sleep spindle detection methods. Artificial Intelligence in Medicine, 2007; 40: 157-170.

Kaida K, Takahashi M, Åkerstedt T, Nakata A, Otsuka Y, Haratani T, Fukasawa K. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical Neurophysiology, 2006; 117: 1574-1581.

Kecklund G, Åkerstedt T. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. Ergonomics, 1993; 36: 1007-1017.

Klimesch W, Sauseng P, Hanslmayr S. EEG alpha oscillations: the inhibition-timing hypothesis. Brain Research Reviews, 2007; 53: 63-88.

Kohlmorgen J, Dornhege G, Braun M, Blankertz B, Müller K-R, Curio G, Hagemann K, Bruns A, Schrauf M, Kincses WE. Improving human performance in a real operating environment through real-time mental workload detection. In: Dornhege G, Millán JR, Hinterberger T, McFarland D, Müller KR (eds). Toward Brain-Computer Interfacing. Cambridge, Mass: MIT press, 2007: 409-422.

Lal SKL, Craig A. A critical review of the psychophysiology of driver fatigue. Biological Psychology, 2001; 55: 173-194.

Lal SKL, Craig A. Driver fatigue: electrophysiology and psychological assessment. Psychophysiology, 2002; 39: 1-9.

Lal SKL, Craig A. Reproducibility of the spectral components of the electroencephalogram during driver fatigue. Psychophysiology, 2005; 55: 137-143.

Lansdown TC. Causes, measures and effects of driver visual workload. In Hancock PA, Desmond PA, editors. Stress, Workload and Fatigue. Lawrence Erlbaum: London, 2001; 351-369.

Lee TW, Girolami M, Sejnowski TJ. Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources. Neural Computation, 1999; 11: 417-441.

Lemke M. Correlation between EEG and driver's actions during prolonged driving under monotonous conditions. Accident Analysis and Prevention, 1982; 14: 7-17.

Lin CT, Wu RC, Jung TP, Liang SF, Huang TY. Estimating Driving Performance Based on EEG Spectrum Analysis. EURASIP Journal on Applied Signal Processing, 2005; 19: 3165-3174.

May JF, Baldwin CL. Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation Research Part F, 2009; 12: 218-224.

Moller HJ, Kayumov L, Bulmash EL, Nhan J, Shapiro CM. Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness. Journal of Psychosomatic Research, 2006; 61: 335-342.

NHTSA Expert Panel. Drowsy Driving and Automobile Crashes. National Highway Traffic Safety Administration, 2003 (http://www.nhtsa.dot.gov/people/injury/drowsy_driving1/Drowsy.html)

O'Brien RG, Kaiser MK. MANOVA Method for Analyzing Repeated Measures Designs: An Extensive Primer. Psychological Bulletin, 1985; 97(2): 316-333.

O'Hanlon JF, Kelly GR. Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In Mackie R, editor. Vigilance. Plenum Press, New York, 1977; 87-109.

Oken BS, Salinsky MC, Elsas SM. Vigilance, alertness, or sustained attention: physiological basis and measurement. Clin Neurophysiol, 2006; 117: 1885-1901.

Oppenheim AV, Schafer RW. Discrete-Time Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1989.

Otmani S, Pebayle T, Roge J, Muzet A. Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Physiology & Behavior, 2005; 84: 715-724.

Papadelis C, Chen Z, Kourtidou-Papadeli C, Bamidis PD, Chouvarda I, Bekiaris E, Maglaveras N. Clin Neurophysiol, 2007; 118: 1906-1922.

Pereda E, Gamundi A, Rial R, Gonzalez J. Non-linear behaviour of human EEG – fractal exponent versus correlation dimension in awake and sleep stages. Neurosci. Lett., 1998; 250: 91-94.

Petsche H, Kaplan S, von Stein A, Filz O. The possible meaning of the upper and lower alpha frequency ranges for cognitive and creative tasks. Int J. Psychophysiology, 1997; 26: 77-97.

Pfurtscheller G, Stancák A, Neuper C. Event-related synchronization (ERS) in the alpha band - An electrophysiological correlate of cortical idling: a review. International Journal of Psychophysiology, 1996; 24: 39-46.

Riemersma JBJ, Sanders AF, Wildervanck C, Gaillard AW. Performance decrement during prolonged night driving. In: Mackie R, editor. Vigilance Theory. Plenum Press, New York, 1977: 41-58.

Ryu SY, Hirata M, Sakihara K, Kimura K, Ebe K, Yoshioka M, Kato A, Yoshimine T, Yorifuji S. Temporal dynamics of wakefulness during simulated driving. International Congress Series, 2007; 1300: 429-432.

Schmidt EA, Schrauf M, Simon M, Fritzsche M, Buchner A, Kincses WE. Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. Accident Analysis and Prevention, 2009; 41: 1087-1093.

Shaw JC. The Brain's Alpha Rhythms and the Mind. New York: Elsevier, 2003.

Shen KQ, Li XP, Ong CJ, Shao SY, Wilder-Smith E. EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate. Clin Neurophysiol, 2008; 119: 1524-1533.

Strijkstra AM, Beersma DG, Drayer B, Halbesma N, Daan S. Subjective sleepiness correlates negatively with global alpha (8–12 Hz) and positively with central frontal theta (4–8 Hz) frequencies in the human resting awake electroencephalogram. Neurosci Lett, 2003; 340(1): 17-20.

Tietze H, Hargutt V. Zweidimensionale Analyse zur Beurteilung des Verlaufs von Ermüdung. 43. Tagung experimentell arbeitender Psychologen, Regensburg, 2001.

Trejo LJ, Knuth K, Prado R, Rosipal R, Kubitz K, Kochavi R et al. EEG-based estimation of mental fatigue: convergent evidence for a three-state model. Lecture Notes Comput Sci 2007; Vol 4565/2007: 201-211.

Van Dijk H, Schoffelen JM, Oostenveld R, Jensen O. Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. Journal of Neuroscience, 2008; 28(8): 1816-1823.

Wagenmakers EJ, Farrell S, Ratcliff R. Estimation and interpretation of $1/f^{\alpha}$ noise in human cognition. Psychonomic Bulletin & Review, 2004; 11 (4): 579-615.

Accident Analysis and Prevention 41 (2009) 1087–1093

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving

Eike A. Schmidt[a,b,*], Michael Schrauf[a,b], Michael Simon[a,c], Martin Fritzsche[a], Axel Buchner[b], Wilhelm E. Kincses[a]

[a] Daimler AG, Research and Development, HPC H602, Leibnizstr. 2, D-71032 Böblingen, Germany
[b] Heinrich-Heine-University Düsseldorf, Institut für Experimentelle Psychologie, Universitätsstr. 1, D-40225 Düsseldorf, Germany
[c] University of Tübingen, Wilhelm-Schickard-Institut für Informatik, Sand 13, D-72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 20 March 2009
Received in revised form 9 June 2009
Accepted 12 June 2009

Keywords:
Vigilance
Driving
Monotony
Self-assessment
Performance
EEG
ECG

ABSTRACT

To investigate the effects of monotonous daytime driving on vigilance state and particularly the ability to judge this state, a real road driving study was conducted. To objectively assess vigilance state, performance (auditory reaction time) and physiological measures (EEG: alpha spindle rate, P3 amplitude; ECG: heart rate) were recorded continuously. Drivers judged sleepiness, attention to the driving task and monotony retrospectively every 20 min. Results showed that prolonged daytime driving under monotonous conditions leads to a continuous reduction in vigilance. Towards the end of the drive, drivers reported a subjectively improved vigilance state, which was contrary to the continued decrease in vigilance as indicated by all performance and physiological measures. These findings indicate a lack of self-assessment abilities after approximately 3 h of continuous monotonous daytime driving.

1. Introduction

Following Mackworth (1957), who defined vigilance as "the state of readiness to detect and respond to certain specified small changes occurring at random time intervals in the environment," driving can be classified as a vigilance task, especially when it is performed in a monotonous environment with little task demand. This also follows Parasuraman's (1998) somewhat broader definition of vigilance as "the ability to sustain attention to a task for a period of time". Fluctuations in vigilance in general and a vigilance decrement in particular constitute a serious risk to traffic safety.

As approximately 15–20% of the fatal accidents can be ascribed to sleepiness and fatigue (Hell and Langwieder, 2001; CARE, EU road accidents database, 2009), Europe was faced with at least 6000 vigilance-related fatalities in 2007, the most common and severe cause of these accidents being the driver falling asleep while driving. Typically carried out in driving simulators, at night and oftentimes with sleep-deprived participants, a fair amount of driving studies investigated the occurrence of microsleeps (Boyle et al., 2008; Moller et al., 2006; Papadelis et al., 2007). However, there is evidence from accident data (Folkard, 1997) as well as from experimental studies (Thiffault and Bergeron, 2003) that vigilance fluctuations have a significant negative impact on driving safety also during daytime driving and especially under monotonous conditions (Dinges, 1995). Unfortunately, there are only few daytime driving studies investigating drivers' vigilance states under monotonous conditions, and only a subset of those studies were conducted in a real road situation (Brookhuis and De Waard, 1993; Tejero and Choliz, 2002). Based on the finding that fatigue develops differently in a simulator as compared to real road driving conditions (Belz et al., 2004; Philip et al., 2005) we decided to conduct a road driving study in real traffic in order to maximize the ecological validity of the results. Given the debate on how well people are able to judge their own vigilance state (see Section 1.2) it seemed appropriate to put a distinct focus on a possible dissociation between the drivers' self-assessment on one side and performance as well as physiological indicators of vigilance on the other.

1.1. Factors affecting vigilance in driving

Thiffault and Bergeron (2003) pointed out that factors affecting vigilance can be divided into exogenous and endogenous factors

* Corresponding author at: Daimler AG, Research and Development, HPC H602, Leibnizstr. 2, 71032 Böblingen, Germany. Tel.: +49 7031 90 60555; fax: +49 711 3052 178179.
E-mail addresses: eike_andreas.schmidt@daimler.com, eikeandreas.schmidt@uni-duesseldorf.de (E.A. Schmidt), michael.schrauf@daimler.com (M. Schrauf), michael.m.simon@daimler.com (M. Simon), martin.fritzsche@daimler.com (M. Fritzsche), axel.buchner@uni-duesseldorf.de (A. Buchner), wilhelm.kincses@daimler.com (W.E. Kincses).

0001-4575/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2009.06.007

depending on whether they stem from within the organism or whether they are caused by characteristics of the task performed. Vigilance can be influenced by monotony which may itself serve as a cause or a multiplier of fatigue or sleepiness (Dinges, 1995; Thiffault and Bergeron, 2003). Monotony of the road environment is therefore classified as an exogenous factor. It constitutes the central factor imposed in our experimental design. In terms of Wertheim's (1991) definition, monotony in driving is not only caused by a lack of alerting stimulation but also by a high predictability of the situation.

Fatigue is exogenously influenced as it follows sustained activity and can be overcome by rest (Philip et al., 2005). Desmond and Hancock (2001) further differentiated between active fatigue following a task that requires continuous and prolonged perceptual-motor adjustments and passive fatigue which develops in a monitoring task with only rare perceptual-motor response requirements. As the task of monotonous driving integrates both sorts of fatigue we prefer the term task-related fatigue here.

Sleepiness can be defined as the difficulty in remaining awake (Philip et al., 2005). It is influenced by circadian and homeostatic variables. It is thus endogenously based and can be reduced after sleep. In our experiment we did not directly manipulate participants' sleepiness (e.g., by sleep deprivation) but focused on introducing a vigilance decrement by task characteristics.

Independent of whether monotony, fatigue or sleepiness (or any combination thereof) reduces drivers' vigilance, the best method to improve a reduced vigilance state is to stop the car and take a break, ideally including a short nap (Horne and Reyner, 1999). Therefore, the drivers' ability to correctly judge their vigilance state is of great importance for traffic safety.

1.2. Self-assessment of vigilance

Research on the reliability of the self-assessment of vigilance state has led to contradicting results. Several studies showed that self-ratings are not sufficiently accurate to serve as reliable and valid indicators of reaction times, driving performance or sleep propensity (Belz et al., 2004; Lenné et al., 1997; Moller et al., 2006; Philip et al., 1997, 2005). For instance, Philip et al. (2003) reported that self-assessment of subsequent performance in a reaction time task was rather poor under prolonged daytime driving conditions.

Data obtained in other studies, most of which were conducted under conditions of sleep deprivation or at night-time, suggested that people are generally capable of judging possible performance impairments in tasks such as in driving (Baranski, 2007; Horne and Baulk, 2004; Lisper et al., 1986; Nordbakke and Sagberg, 2007). Therefore one could argue that people are well aware of their deteriorating vigilance, but that early warning signs are often ignored or misinterpreted. Another explanation for these disputed results might have been the use of different vigilance measures across the different studies. In order to gain a more complete picture of the drivers' self-assessment ability, we decided to study a set of objective and subjective measures in a single study.

1.3. Objective measures and their use in studies on monotonous driving

1.3.1. Performance measures

Given that a reduction in vigilance is actually defined by a performance decrement (Mackworth, 1957), performance measures have high face validity for the evaluation of vigilance states. From a safety perspective we had to avoid a state where ongoing driving performance would be too seriously degraded. Nevertheless it was our objective to detect even minor changes in vigilance state that might lead to a reduced ability to respond to unforeseen events. We therefore decided to implement a simple secondary auditory task that most likely would not interfere with the motor requirements of the driving task, assuming that even minor reductions in vigilance should first be reflected in secondary task performance. To ensure that the auditory task would be performed as a secondary task in accordance with the subsidiary task paradigm (O'Donnell and Eggemeier, 1986), participants were explicitly instructed to prioritize the primary task of driving. In addition, it was to be expected that the potentially high costs of errors would also cause participants to give the highest possible priority to the driving task.

In the context of a real driving situation, Laurell and Lisper (1978) demonstrated that a secondary task was sensitive to changes in vigilance and predictive of brake reaction time. Other studies showed that slow reactions (as opposed to fast ones) are particularly sensitive indicators of states of reduced vigilance (Graw et al., 2004; Williams et al., 1959).

1.3.2. Physiological measures

As the primary physiological method we used electroencephalography (EEG) to measure the driver's brain activity continuously. From the various measures that can be derived from the EEG we decided to use the spontaneous alpha spindle rate that is a feature derived from the alpha-band (6–13 Hz) which has been shown to correlate with changes in vigilance in the driving context (Kecklund and Akerstedt, 1993; Papadelis et al., 2007; Tietze and Hargutt, 2001). Based on these earlier findings we implemented an automated algorithm that extracts the alpha spindle rate from the continuous EEG. The main reasons for preferring this measure to the classic power measures are its robustness against external noise and artifacts as well as its superior specificity to changes in vigilance. We also assessed the amplitude of the stimulus-induced P3 event related potential (ERP, for a review see Polich, 2007) that has also been shown to be sensitive to changes in vigilance (Koelega et al., 1992). The P3 amplitude can be interpreted as a measure of the processing depth of the auditory stimulus. Additionally, the participant's heart rate was recorded as an indicator of the physical activation level which has also shown to be sensitive to vigilance changes (O'Hanlon and Kelly, 1977). For a thorough review of objective driver state measures we refer to Tejero Gimeno et al. (2006).

1.4. Hypothesis

We hypothesized that all objective measures would indicate a monotonous reduction in vigilance state as a function of the distance driven (i.e. an increase of the mean of the slow reaction times; an increase of the alpha spindle rate; a decrease of the P3 amplitude and heart rate). Following the controversy concerning the drivers' self-assessment ability we refrained from formulating a specific hypothesis about the effect of driving distance on subjective measures of vigilance.

2. Methods

2.1. Participants

Twenty-nine right-handed participants (20 males, 9 females; age: M = 29.2; range: 23–49) with extensive driving experience (mean driving distance of approximately 20,500 km, i.e. 12,000 miles per year) were recruited on a voluntary basis for an "in-car EEG-study on attentional processes". Participants were screened for a variety of exclusion criteria (handedness, auditory and visual disabilities, and various illnesses), instructed to sleep regularly the night before the experiment, and to refrain from consuming caffeine in the morning on the day of the experiment. For their participation they received compensation in form of a gift worth approximately € 25.

The size of the sample available for data analysis was reduced due to technical problems leading to insufficient EEG data quality

Table 1
Subjective measures.

Concerning the time period since the last prompting. . .									
KSS	. . .how would you describe your predominant state?								
	Extremely alert		Alert		Neither alert nor sleepy		Sleepy, but no difficulty remaining awake		Extremely sleepy, fighting sleep
	1	2	3	4	5	6	7	8	9
ATT	. . .how attentively have you been driving?								
	Extremely attentively		Attentively		Neither attentively nor inattentively		Inattentively		Extremely inattentively
	1	2	3	4	5	6	7	8	9
MON	. . .how did you perceive the drive?								
	Extremely varied		Varied		Neither varied nor monotonous		Monotonous		Extremely monotonous
	1	2	3	4	5	6	7	8	9

(six participants), lack of compliance (two participants) and fatigue-related break-offs (three participants). As a result, 19 (14 males, 5 females; age: $M = 29.4$, range: 23–49) complete data sets containing all measures were available for statistical analysis.

2.2. Materials and procedures

On the test day the participants arrived at the lab at 10:30 am and signed an informed consent form. While the physiological recording equipment was applied, the participants completed German versions of the morningness–eveningness–questionnaire (D-MEQ; Griefahn et al., 2001) and the Edinburgh handedness inventory (Oldfield, 1971). To control for possible circadian (Folkard, 1997; Lenné et al., 1997) and nutritional effects (Smith and Miles, 1986) all participants had lunch at 11:30 am before a 30-min EEG-baseline containing typical body movements and resting-EEG was recorded in the car. The baseline was recorded in order to allow for research on advanced EEG-artifact processing; these data are not presented here.

Following the German legal maximum driving duration for commercial drivers (4.5 h) the length of the drive was set to approximately 4 h, resulting in an experimental course of 428 km (about 267 miles) on the A 81 Autobahn (between exits Ehningen and Gottmadingen). This route was subdivided into four sections of 107 km length, each section thus corresponding to about an hour of driving time. In an attempt to ensure as much monotony as possible, the participants drove on this low-traffic highway off rush hour times. For practical reasons we started and ended the experiment close to an urban area, which implies that the initial and final sections comprised several kilometres on an interurban Autobahn. According to statistics of the German Federal Highway Research Institute (Fitschen et al., 2007) this resulted in a significantly higher average traffic density for sections 1 and 4 as compared to sections 2 and 3 (Fig. 1).

Participants started the drive at 12:45 pm and returned, on average, after about 3:45 h of driving, except for cases in which the experiment was terminated early by the participants. Three predefined turns were necessary and interrupted the continuous run at about 1:00, 1:40, and 2:20 h cumulated driving duration.

Participants had to be sufficiently rested upon arrival. They knew that they could stop driving at any time without any monetary or other penalties. This occurred in three cases. For additional safety reasons, an investigator accompanied the participant in the car, continuously monitoring the driver and ready to intervene whenever necessary. The test car was a Mercedes-Benz S-Class (W221). The participants' task was to drive at a speed not exceeding 130 km/h (approximately 80 mph; recommended maximum speed on German highways) and to comply with the traffic rules at all times. They were instructed to use automatic shift and to refrain from turning on the radio or using other in-car devices. Further, participants were asked not to talk to the investigator and to avoid unnecessary movements in order to reduce artifacts in the EEG recording.

2.3. Subjective measures

As a compromise between a high temporal resolution and a low amount of intrusion we decided to prompt the drivers every 20 min for their retrospective vigilance assessment. As vigilance indicators we used a well-established single-item indicator of sleepiness (Karolinska Sleepiness Scale, KSS; Åkerstedt and Gillberg, 1990) and two similarly constructed items assessing inattention (ATT) and monotony (MON, for an overview see Table 1). The investigator verbally prompted the driver to judge sleepiness, inattention and monotony with regard to the previous 20 min of driving time.

2.4. Performance measures

The participants were instructed that the auditory oddball reaction time task was only to be completed if they felt that it was safe to do so in a given driving situation. The participants had to respond to infrequent target tones (500 Hz, 20% probability) that were presented in a random sequence mixed with frequent distractor tones (400 Hz, 80%) by pressing a button fitted to their right thumb. The inter-stimulus interval varied randomly between 4 s and 6 s. The

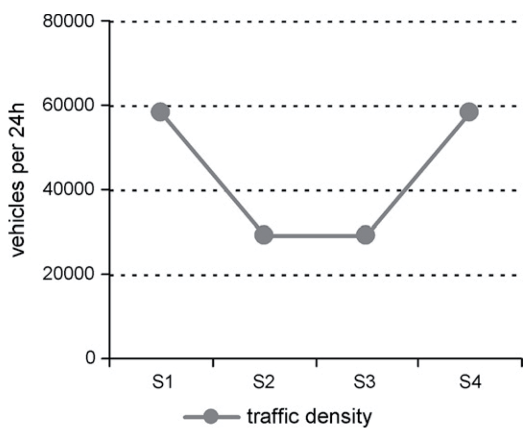

Fig. 1. Average 24 h traffic density.

constant pitch of the stimuli reduced their alerting potential to a minimum. Only button presses that fell into a response window of 200–4000 ms following stimulus presentation were analyzed. For every route section of 107 km, and separately for each participant, the mean of all reaction times above the 80%-percentile was calculated as a measure of the participants' slow reactions, and the mean of all reaction times below the 20%-percentile was calculated as a measure of participants' fast reactions. Consistent with Williams et al. (1959) we decided to focus on the mean of the slowest reactions as these are known to be particularly sensitive to changes in vigilance state.

2.5. Physiological measures

EEG and electrocardiogram (ECG) were recorded from 128 electrodes (1000 Hz sampling rate, low cut-off: 0.016 Hz; high cut-off: 250 Hz) using BrainAmp recording hardware (Brainproducts GmbH, Munich). The EEG signal was down-sampled to 250 Hz and band-pass filtered (0.5–50 Hz); artifactual channels were excluded from further analysis. In order to minimize ocular and muscular artifacts, independent component analysis (Jung et al., 2000) was applied. Only those components carrying a temporal and spatial pattern resembling that of neural sources were accepted.

We used an automated algorithm to extract sharp spectral peaks within the alpha band (6–13 Hz), which we call alpha spindles, and to determine amplitude, peak frequency and duration of these peaks. In this paper we focus on the alpha spindle rate, which is the occurrence rate of alpha spindles within each of the four sections of the drive. As the alpha rhythm is most prominent over parieto-occipital sites we analyzed the signal of electrode Pz. To account for inter-individual differences an alpha spindle index was derived by dividing each section by the reference value of the first section.

To extract the P3 amplitudes the pre-processed EEG signal of electrode CPz was averaged time locked to the presentation of the oddball stimulus. A baseline-correction (relative to −200 ms to 0 ms pre-stimulus time window) was applied. The P3 amplitude then was defined as the maximum value of the signal in a time window from 300 ms to 600 ms post-stimulus minus the minimum value in a time window from 0 ms to 300 ms post-stimulus.

R-peaks were identified from the ECG using an automated algorithm in Matlab and the average heart rate was calculated for every experimental block.

2.6. Data reduction

In order to assess the subjective, performance and physiological correlates of monotonous driving, data epochs which clearly lacked monotonous driving (i.e. communication between driver and investigator, turning points, workload-inducing driving situations, traffic jams and short stops) were discarded from the data analysis. For that purpose, all situations which were noticeably non-monotonous were logged by the investigator accompanying the participant. This resulted in an average loss of 11.9% of the data (SD = 2.9). For all measures the mean over each section of 107 km length was calculated.

2.7. Experimental design

The only independent variable in our experimental design was the distance driven (four sections of 107 km each). Dependent variables were (a) subjective measures of vigilance (sleepiness [KSS], inattention [ATT] and monotony [MON]), (b) performance measures (slow and fast reactions) and (c) physiological measures of vigilance (alpha spindle rate (ASR), P3 amplitude (P3A) and heart rate (HR)).

An a priori statistical power analysis using G*Power 3 (Faul et al., 2007) showed that in order to detect effects of $f = .50$ assuming a population correlation among the levels of the repeated measures variable of $\rho = .50$ (estimated from pilot data) and given $\alpha = \beta = .05$, $N = 13$ participants were needed. A sensitivity power analysis showed that given a final sample of $N = 19$, effects of size $f = .39$ could be detected under otherwise identical conditions. A multivariate approach (MANOVA) was used for all within-subject comparisons to identify the effect of the driving distance variable for each dependent measure. All multivariate test criteria correspond to the same (exact) F-statistic, which is reported. The level of α was set to .05 for all analysis. Whenever H_0 had to be rejected, the partial η^2 is reported as a measure of relative effect size. Statistically significant results were subjected to post hoc trend-analyses using polynomial contrasts. Only significant linear and quadratic trends are reported.

3. Results

It was confirmed that all participants were right-handers (all handedness-indices >0). The D-MEQ results showed that the large majority of participants (15) fell into the neutral chronotype group. Considering the small variability in chronotype and the small sample size, we refrained from entering chronotype as a further factor into the analysis.

The results of the MANOVAs testing the effect of the driving distance variable (sections 1–4) on the dependent variables are summarized in Table 2. Significant quadratic trends for KSS and ATT indicate a relatively better subjective state at the beginning and at the end of the drive compared to the two middle sections (Fig. 2). The same pattern of results was found for the measure MON although the main effect barely failed to reach significance (Fig. 2).

Table 2
Statistical results.

	Main effect			Trend analysis			
	$F(3,15)$	p	η^2	Type	$F(1,17)$	p	η^2
Subjective							
MON	3.23	.053	.392	Quadratic	9.19	.008	.351
KSS	8.49	.002	.629	Quadratic	9.99	.006	.370
ATT	4.47	.020	.472	Quadratic	10.06	.006	.372
Performance							
Fast reactions	1.74	n.s.					
Slow reactions	6.38	.005	.561	Linear	13.36	.002	.440
Physiology							
ASR	5.05	.013	.503	Linear	9.77	.006	.365
P3A	3.74	.035	.428	Linear	10.25	.005	.376
Heart rate	11.86	<.001	.703	Linear	17.83	.001	.512

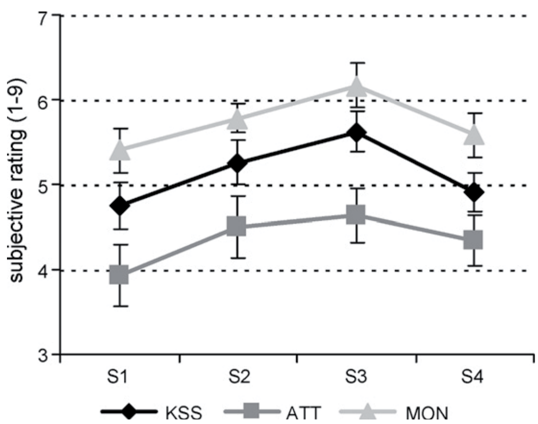

Fig. 2. Subjective measures (rating: 1–9). Error bars represent the standard errors of the means.

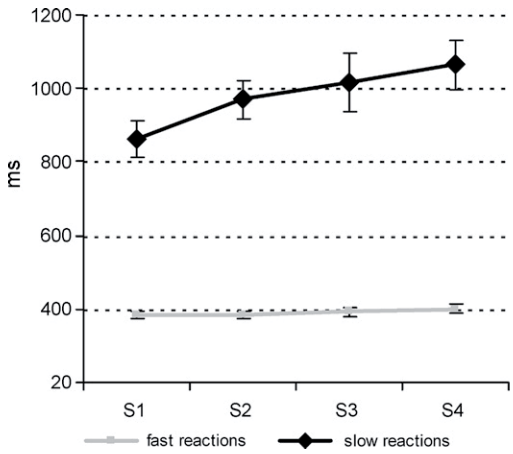

Fig. 3. Reaction times. Error bars represent the standard errors of the means.

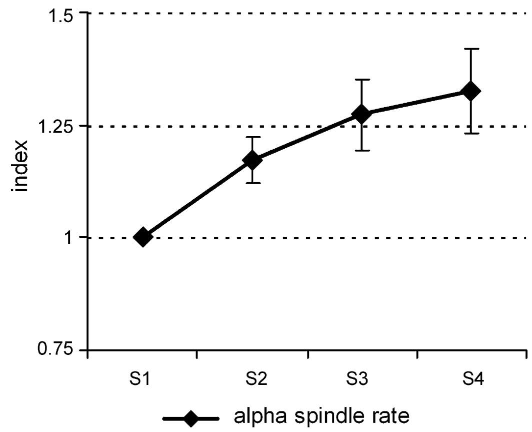

Fig. 4. Alpha spindle rate at electrode Pz (relative to first section). Error bars represent the standard errors of the means.

As expected, the mean of the fast reactions did not differ significantly among the four sections. In contrast, the mean of the slow reactions showed a significant effect in terms of a linear increase from section 1 to 4 (Fig. 3).

The same pattern of results was observed for the physiological measures (alpha spindle rate, P3 amplitude and heart rate, see Figs. 4–6). There was a linear increase in the alpha spindle rate and a linear decrease in the P3 amplitude and as well as the heart rate.

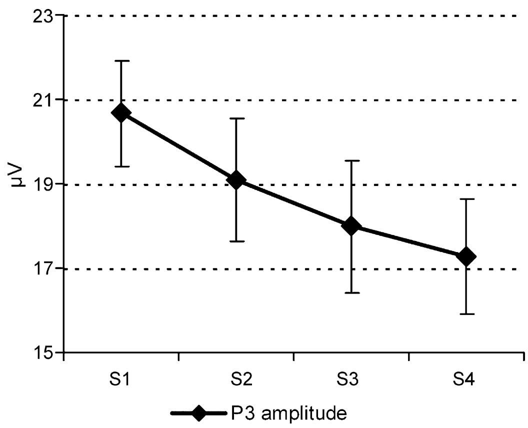

Fig. 5. P3 amplitude at electrode CPz. Error bars represent the standard errors of the means.

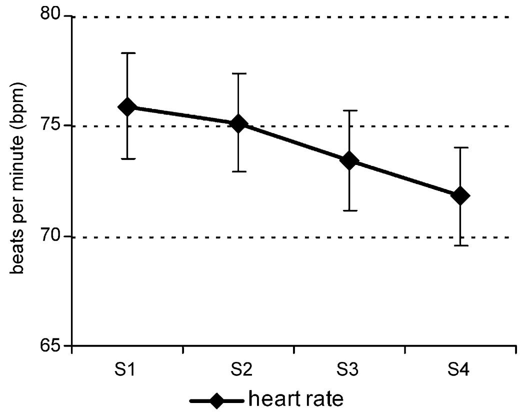

Fig. 6. Heart rate. Error bars represent the standard errors of the means.

4. Discussion

The participants' subjective evaluation of their vigilance level indicates that the induction of monotony was successful. It is also obvious that the quasi-experimental variation in traffic density led to a reduction of monotony at the beginning and at the end of the drive. Most interestingly, drivers reported a subjective improvement of their vigilance (as indicated by their assessments of their sleepiness and attention for the driving task) towards the end of the drive. However, this subjective improvement contradicted all objective measures of vigilance. Reaction times, EEG alpha spindle rate, P3 amplitude and heart rate consistently indicated a continuous reduction in vigilance even for the fourth and final section of the drive. This leads us to conclude that drivers in a reduced state of vigilance following prolonged and monotonous driving are vulnerable to a misjudgement of their objective vigilance state in terms of performance and physiological parameters.

4.1. Reason for misconception

It is not possible to clearly identify whether the subjective increase in vigilance in the final section of the drive was due to effects of circadian phase, increased traffic density, the (joyful) expectation that the drive would soon be over, or any combination of these variables. All these variables were confounded in the final section of the drive. This state of affairs does not reduce the practical relevance of the finding that self-assessment of vigilance may dissociate from objective measures of vigilance. Further research is needed to shed light on the interrelations of the variables that may cause this potentially fatal dissociation.

4.2. Reaction times

The increase in the slow reactions as compared to the fast reactions points to the validity of the separate analysis of the long reaction times. A possible problem is that the increase in traffic density in the fourth section of the drive may have resulted in a shift of attentional capacity away from the reaction time task towards the driving task. If so, then the observed increase in the long reaction times could be the result of an increased workload rather than a decrease in vigilance. However, this alternative explanation seems very unlikely, because a shift of attention towards the visual modality would have to be accompanied by a reduction in alpha spindle rate. This was clearly not the case. Conversely, if one wanted to interpret the EEG in light of an attentional shift, the observed increase in alpha spindle rate would imply a switch of attention towards the auditory task (Gladwin and de Jong, 2005). Further an increase in workload in the last section of the drive would have to be accompanied by an increase in heart rate. This was not observed either. Finally, given that the traffic density in sections 1 and 4 was nearly identical, reaction times also should have been nearly identical in these two sections if the speed of responding simply reflected traffic density related workload differences. Obviously, this was not the case.

While we cannot conclude with certainty from our data that the deterioration of reaction time performance in our auditory secondary task goes along with a reduced driving ability in terms of a reduced speed in responding to unforeseen events, it is possible to point to earlier research showing a positive correlation between auditory reaction times (Laurell and Lisper, 1978). In our case, travelling at a speed of 130 km/h the observed increase in slow reaction times of 200 ms in average corresponds to a potential increase in stopping distance of approximately 7 m. This significant increase underscores the safety relevance of our findings.

4.3. Future research

The present study was designed to investigate one driving course performed at a fixed time of day. Future studies should be aimed at uncovering how well the pattern of results observed here generalizes to other road environments and to different times of day and night. Future research is also needed to identify and evaluate potential countermeasures that, on one hand, decrease driver's proneness to the effects of monotonous driving and, on the other hand, support the driver's ability to judge his or her own state correctly. The present data show that this judgement ability is vulnerable under certain conditions. Research is thus needed on the acceptance of modern driver monitoring systems. Such systems might correctly judge the driver's state as inattentive while the driver feels alert enough to continue his drive. Educative actions communicating the shortcomings of human self-monitoring might be a promising approach here. Finally, performance and physiological measures might be suited to objectively identify particularly monotonous and therefore dangerous road environments that lead to an accelerated vigilance decrement, but this assumption, too, needs to be validated experimentally.

5. Conclusion

In conclusion, the data presented here show a good correspondence between subjective measures and reaction times as well as physiological measures for the first three of four sections of driving a long distance. All measures indicated a decrease in vigilance. However, in the fourth and final section of the drive a clear dissociation was observed between subjective and objective measures of vigilance. Given that all objective measures consistently point to a further decrease in vigilance in the fourth section, we must conclude that participants misjudged their subjective states at the end of the drive. Factors such as a less monotonous driving situation in the last section, the expectancy that the trip will be over soon or circadian phase might all have contributed to this misjudgement.

The present findings further support the potential benefit of driver assistance systems that constantly monitor the driver's state. This is especially true for systems that are sensitive to the early levels of inattentive driving. Given appropriate acceptance, these systems supply drivers with an objective evaluation of their ability to drive, they can keep drivers from continuing to drive following a misjudgement of their own state, and therefore they can reduce the probability of vigilance-related accidents.

Acknowledgements

This research was supported by BMBF (German Federal Ministry of Education and Research) Grant 16SV2233. Parts of this research were presented at the Fourth International Symposium on Human Factors in Driving Assessment, Training, and Vehicle Design (Stevenson, Washington, July 9–12, 2007). We thank Prof. Gabriel Curio and Dr. Ruth Schubert for valuable discussions on the experimental design; Stefan Haufe, Andreas Pröttel, Sven Willmann and Claus Aufmuth for support with data recording and analysis; Jasmin Steinwender and Hendrik von Hülst for their help with data collection; Melinda Ewert-Kincses for critical reading.

References

Åkerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. International Journal of Neuroscience 52, 29–37.
Baranski, J.V., 2007. Fatigue, sleep loss, and confidence in judgment. Journal of Experimental Psychology: Applied 13 (4), 182–196.
Belz, S.M., Robinson, G.S., Casali, J.G., 2004. Temporal separation and self rating of alertness as indicators of driver fatigue in commercial motor vehicle operators. Human Factors 46 (1), 154–169.
Boyle, L.N., Tippin, J., Paul, A., Rizzo, M., 2008. Driver performance in the moments surrounding a microsleep. Transportation Research Part F 11, 126–136.
Brookhuis, K., De Waard, D., 1993. The use of psychophysiology to assess driver status. Ergonomics 36 (9), 1099–1110.
CARE, road accidents database, 2009. Road Safety Evolution in EU., http://ec.europa.eu/transport/road_safety/observatory/doc/historical_evol.pdf.
Desmond, P.A., Hancock, P.A., 2001. Active and passive fatigue states. In: Hancock, P.A., Desmond, P.A. (Eds.), Stress, Workload, and Fatigue. LEA Publishers, London, pp. 455–465.
Dinges, D.F., 1995. An overview of sleepiness and accidents. Journal of Sleep Research 4, 4–14.
Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 39, 175–191.
Fitschen, A., Koßmann, I., Bundesanstalt für Straßenwesen, 2007. BASt-Bericht V 160: Traffic Development on Federal Trunk Roads in 2005.
Folkard, S., 1997. Black times: temporal determinants of transport safety. Accident Analysis and Prevention 29 (4), 417–430.
Gladwin, T.E., de Jong, R., 2005. Bursts of occipital theta and alpha amplitude preceding alternation and repetition trials in a task-switching experiment. Biological Psychology 68, 309–329.
Graw, P., Kräuchi, K., Knoblauch, V., Wirz-Justice, A., Cajochen, C., 2004. Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. Physiology and Behavior 80, 695–701.
Griefahn, B., Künemund, C., Bröde, P., Mehnert, P., 2001. Zur Validität der deutschen Übersetzung des Morningness-Eveningness-Questionnaires von Horne und Östberg. Somnologie 5, 71–80.
Hell, W., Langwieder K., 2001. Einschlafunfälle im Straßenverkehr–Eine bisher oft verkannte Unfallursache. Gesamtverband der Deutschen Versicherungswirtschaft.
Horne, J., Reyner, L., 1999. Vehicle accidents related to sleep: a review. Occupational and Environmental Medicine 56, 289–294.
Horne, J.A., Baulk, S.D., 2004. Awareness of sleepiness when driving. Psychophysiology 4, 97–110.
Jung, T.P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163–178.
Kecklund, G., Akerstedt, T., 1993. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. Ergonomics 36 (9), 1007–10017.
Koelega, H.S., Verbaten, M.N., van Leeuwen, T.H., Kenemans, J.L., Kemner, C., Sjouw, W., 1992. Time effects on event-related brain potentials and vigilance performance. Biological Psychology 34, 59–86.
Laurell, H., Lisper, H.O., 1978. A validation of subsidiary reaction time against detection of roadside obstacles during prolonged driving. Ergonomics 21, 81–88.
Lenné, M.G., Triggs, T.J., Redman, J.R., 1997. Time of day variations in driving performance. Accident Analysis and Prevention 29 (4), 431–437.
Lisper, H.O., Laurell, H., van Loon, J., 1986. Relation between time to falling asleep behind the wheel on a closed track and changes in subsidiary reaction time during prolonged driving on a motorway. Ergonomics 29 (3), 445–453.
Mackworth, N.H., 1957. Vigilance. The Advancement of Science 53, 389–393.
Moller, H.J., Kayumov, L., Bulmash, E.L., Nhan, J., Shapiro, C.M., 2006. Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness. Journal of Psychosomatic Research 61, 335–342.
Nordbakke, S., Sagberg, F., 2007. Sleepy at the wheel: knowledge, symptoms and behaviour among car drivers. Transportation Research Part F 10, 1–10.
O'Donnell, R.D., Eggemeier, F.T., 1986. Workload assessment methodology. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.), Handbook of Perception and Human Performance. Volume II, Cognitive Processes and Performance. Wiley, New York, pp. 42/1–142/.
O'Hanlon, J.F., Kelly, G.R., 1977. Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In: Mackie, R. (Ed.), Vigilance. Plenum Press, New York, pp. 87–109.
Oldfield, R.C., 1971. The assessment and analysis of handedness. The Edinburgh inventory. Neuropsychologia 9, 97–113.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P.D., Chouvarda, I., Bekiaris, E., Maglaveras, N., 2007. Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. Clinical Neurophysiology 118, 1906–1922.
Parasuraman, R., 1998. In: Parasuraman, R. (Ed.), The Attentive Brain. The MIT Press, Cambridge, MA.
Philip, P., Ghorayeb, I., Leger, D., Menny, J.C., Bioulac, B., Dabadie, P., Guilleminault, C., 1997. Objective measurement of sleepiness in summer vacation long-distance drivers. EEG Journal of Clinical Neurophysiology 102, 383–389.
Philip, P., Sagaspe, P., Taillard, J., Guilleminault, C., Sanchez-Ortuno, M., Åkerstedt, T., Bioulac, B., 2003. Fatigue, sleep restriction, and performance in automobile drivers: a controlled study in a natural environment. Sleep 26 (3), 277–280.
Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Åkerstedt, T., Charles, A., Bloulac, B., 2005. Fatigue, sleepiness, and performance in simulated versus real driving conditions. Sleep 28 (12), 1511–1516.
Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clinical Neurophysiology 118 (10), 2128–2148.
Smith, A.P., Miles, C., 1986. The effects of lunch on cognitive vigilance tasks. Ergonomics 29 (10), 1251–1261.
Tejero, P., Choliz, M., 2002. Driving on the motorway: the effect of alternating speed on driver's activation level and mental effort. Ergonomics 45 (9), 605–618.
Tejero Gimeno, P., Pastor Cerezuela, G., Chóliz Montañés, M., 2006. On the concept and measurement of driver drowsiness, fatigue and inattention: implications for countermeasures. International Journal of Vehicle Design 42, 67–86.
Thiffault, P., Bergeron, J., 2003. Monotony of road environment and driver fatigue: a simulator study. Accident Analysis and Prevention 35, 381–391.
Tietze, H., Hargutt, V., 2001. Zweidimensionale Analyse zur Beurteilung des Verlaufs von Ermüdung. 43. Tagung experimentell arbeitender Psychologen, Regensburg., http://www.psychologie.uni-wuerzburg.de/methoden/texte/2001_tietze_hargutt_Zweidimensionale_Analyse.pdf.
Wertheim, A.H., 1991. Highway hypnosis: a theoretical analysis. In: Gale, A.D., et al. (Eds.), Vision in Vehicle III. Elsevier, North-Holland, pp. 467–472.
Williams, H.L., Lubin, A., Goodnow, J.J., 1959. Impaired performance with acute sleep loss. Psychological Monographs: General and Applied 73 (14), 1–26.

The short-term effect of verbally assessing drivers' state on vigilance indices during monotonous daytime driving

Corresponding author:

Eike A. Schmidt[a] *eikeandreas.schmidt@uni-duesseldorf.de*

phone: +49 176 2060 2075; fax: +49 211 811 5037

Michael Schrauf[a,b] *michael.schrauf@daimler.com*

Michael Simon[b,c] *michael.m.simon@daimler.com*

Axel Buchner[a] *axel.buchner@uni-duesseldorf.de*

Wilhelm E. Kincses[b] *wilhelm.kincses@daimler.com*

[a] Heinrich-Heine-University Düsseldorf; Institut für Experimentelle Psychologie; Universitätsstr. 1; D-40225 Düsseldorf

[b] Daimler AG; Research and Development; GR/PAP; HPC 50-G024; D-71059 Sindelfingen

[c] University of Tübingen; Wilhelm-Schickard-Institut für Informatik; Sand 13; D-72076 Tübingen

Abstract

To investigate the effects of verbal assessment of subjective driver state on objective indicators of vigilance state during a monotonous daytime drive, a real road driving study was conducted. During a four-hour drive participants' subjective state (sleepiness, inattention, monotony) was assessed every 20 minutes by an investigator accompanying the drive. The assessment procedure consisted of roughly one minute of verbal interaction. Physiological indicators (EEG alpha spindle rate, blink duration, heart rate) revealed a significant improvement of vigilance state during the communication episode as compared to a pre-assessment baseline. The activation persisted for up to two minutes following the end of the verbal interaction. Reaction times supported these findings by indicating a significant decrease after the communication. The P3 amplitude of the auditory event related potential did not show any consistent results. It can be concluded that a short verbal assessment has positive effects on drivers' vigilance state. However, these effects persist only for a very limited time. The implications of these findings for the frequency of verbal assessment during experimental studies and for the use of verbal communication as a fatigue countermeasure are discussed.

Keywords: Vigilance; Fatigue; Monotony; Driving; Self assessment; Fatigue countermeasures

1. Introduction

The negative effects of prolonged monotonous night- and daytime driving on driver vigilance have been repeatedly investigated (e.g., Brookhuis & De Waard, 1993; Horne & Reyner, 1999; O'Hanlon & Kelly, 1978; Philip, Sagaspe, Taillard, Valtat, Moore, Åkerstedt et al., 2005; Schmidt, Schrauf, Simon, Fritzsche, Buchner, & Kincses, 2009; Thiffault & Bergeron, 2003). In an attempt to reduce the number of fatigue-related accidents, numerous researchers have been working on the task of reliably detecting vigilance-related deficits in drivers (for reviews see Kecklund, Åkerstedt, Sandberg, Wahde, Dukic, Anund et al., 2006; Wright, Stone, Horberry, & Reed, 2007). The general absence of a single valid measure for every individual (Kecklund et al., 2006) makes it necessary to acquire a wide variety of vigilance indicators including subjective self-assessment measures of driver state. Subjective measures, often assessed verbally, bear the potential of influencing the state being under investigation (Kaida, Åkerstedt, Kecklund, Nilsson, & Axelsson, 2007) and by their nature cannot be recorded continuously. The objective of this study was to evaluate the magnitude and especially the duration of the potential influence of verbal assessment on driver state in real traffic. This should indicate how often a subjective assessment is feasible without significantly influencing the state under investigation. As the verbal assessment resembles a short communication between driver and passenger, the presented findings also allow inferring its potential as a countermeasure to a vigilance decrement under monotonous daytime driving conditions.

1.1 Terminology: Fatigue and Sleepiness

Sleepiness and fatigue are often used synonymously. For clarification May and Baldwin (2009) introduced their model of fatigue, distinguishing between task-related active as well as passive fatigue and sleep-related fatigue, which most authors refer to as sleepiness or the difficulty in remaining awake (Philip et al., 2005). May and Baldwin state that the performance decrement - and therefore the task of detecting it - might be similar for all types of fatigue whereas countermeasures might only work for one or the other. Given that we investigated monotonous daytime driving with well-rested drivers, our focus is on task-related passive fatigue, although strictly speaking a minor influence of sleep-related fatigue could not be ruled out because our participants partly drove during their circadian afternoon dip.

1.2. Subjective Assessment of Driver State

Various single- and multi-item questionnaire measures of sleepiness have been developed and have been validated against both physiological as well as performance measures (e.g., Stanford Sleepiness Scale “SSS”: Hoddes, Zarcone, Smythe, Phillips, & Dement, 1973; Karolinska Sleepiness Scale “KSS”: Åkerstedt & Gillberg, 1990; Kaida, Takahashi, Åkerstedt, Nakata, Otsuka, Haratani et al., 2006). Apart from their limited objectivity, subjective questionnaire measures of sleepiness imply a tradeoff between the number of measurements in a given time period and the possibility of interfering with the process under investigation because answering questions may have activating effects and may thus reduce sleepiness. According to Kecklund et al. (2006) the measurement frequency of the KSS varies from once every three to five minutes down to just a few assessments per hour. Kaida et al. (2007) found that indeed the repeated rating of sleepiness reduced subjective post-test sleepiness, reduced EEG-alpha power, and improved the subjective perception of performance. Interestingly, task performance did not improve. This led the authors to conclude that the effects of sleepiness were underestimated in the sleepiness ratings. Kaida et al. were interested in the long-term effects of sleepiness ratings for an entire experimental session. Therefore, conclusions about the immediate influence of the verbal sleepiness assessments were not obtained.

1.2.1 Indications from Studies on Countermeasures

Although little is known about the short-term effects of a subjective state assessment in a monotonous situation, related research on fatigue countermeasures might shed a light on this issue. A variety of experimental studies systematically evaluating the effects of potential countermeasures on sleep-related fatigue have been conducted by Reyner and Horne (e.g., 1997; 1998). They concluded that the most effective and enduring countermeasure to sleepiness was a combination of caffeine intake and a short nap. In comparison, any positive effects of exposure to cold air and turning on the radio lasted only for about 15 minutes. However, questionnaire-based studies showed that in retrospective assessments between 25% and 35% of the drivers considered the engaging in a conversation with a passenger to be useful (Anund, Kecklund, Peters, & Åkerstedt, 2008; Maycock, 1997; Nordbakke & Sagberg, 2007). Therefore, at least some positive effects of investigator-participant communications are to be expected, if only for a brief post-communication interval.

1.2.2 Effect of Conversations

To our knowledge the potentially activating effects of a verbal communication under monotonous driving conditions have been investigated in simulator studies only (Chan & Atchley, 2009; Drory, 1985; Gershon, Ronen, Oron-Gilad, & Shinar, 2009; Oron-Gilad, Ronen, & Shinar, 2008). Drory reported that long-haul truck drivers' performance improved during a monotonous simulator drive when they were engaged in a short verbal task from time to time. Oron-Gilad et al. and Gershon et al. showed that a trivia game, carrying a verbal component, prevented deterioration of simulator driving performance and improved alertness assessed by physiological measures. This effect occurred immediately when engaged in the task but did not last longer than the task itself.

Despite this, studies on the distracting effects of cell phone and passenger conversations (Caird, Willness, Steel, & Scialfa, 2008; Drews, Pasupathi, & Strayer, 2008; Horrey &

Wickens, 2006) have suggested that distraction might outweigh any positive effect of these types of conversations on driving performance. It is beyond the scope of this paper to disentangle these effects. We mainly focus on the effect *following* the communication.

1.3 Assessing Drivers' Vigilance State

The vigilance measures applied in the research reported here can be classified as follows:

- Continuous physiological measures with high temporal resolution and the possibility to take measurements both during and after the investigator-driver communication.
- Discrete event-related measures elicited by an auditory stimulus (providing reaction times at the psychophysical, and event-related brain potentials at the neurophysiological side) that (a) need a sufficient number of stimuli for a reliable assessment and (b) cannot reasonably be recorded during the communication sequence.

1.3.1 Continuous Measures

In accordance with our previous research we used the *alpha spindle rate*, a feature derived from the EEG alpha-band (7 - 13 Hz), as a sensitive EEG-based measure of a vigilance decrement induced by a monotonous driving situation (Schmidt et al., 2009; see also Kecklund & Åkerstedt, 1993; Papadelis, Chen, Kourtidou-Papadeli, Bamidis, Chouvarda, Bekiaris et al., 2007; Tietze & Hargutt, 2001). Apart from its high specificity to changes in vigilance, the main reason for preferring this measure to the classic EEG power measures is its robustness against external and internal noise and artifacts.

We also measured eye-blink duration because several authors reported *blink duration* to be their most informative occulomotoric parameter when measuring fatigue (Caffier, Erdmann, & Ullsperger, 2003; Ingre, Åkerstedt, Peters, Anund, & Kecklund, 2006). In a large sample study Schleicher, Galley, Briest, and Galley (2008) observed a gradual increase in blink duration with decreasing alertness.

Finally participants' *heart rate* was recorded as an indicator of the physical activation level, which has shown to be a sensitive indicator of vigilance changes in the driving context (O'Hanlon & Kelly, 1977).

1.3.2 Discrete Event-Related Measures

In the context of a real driving situation, Laurell and Lisper (1978) demonstrated that an auditory secondary reaction time task was sensitive to changes in vigilance, and that it predicted brake reaction times. *Slow reactions* (as opposed to fast ones) seem to be particularly sensitive indicators of reduced vigilance (Graw, Kräuchi, Knoblauch, Wirz-Justice, & Cajochen, 2004; Williams, Lubin, & Goodnow, 1959) even in the driving context (Schmidt et al., 2009). Accordingly we implemented a simple secondary auditory task that most likely would not interfere with the motor requirements of the driving task, assuming that even minor reductions in vigilance should first be reflected in secondary task performance (subsidiary task paradigm: O'Donnell & Eggemeier, 1986). Participants were explicitly instructed to prioritize the primary task of driving and it was to be expected that the potentially high costs of major driving errors would also cause participants to give the highest possible priority to the driving task.

We further assessed the *amplitude* of the stimulus-induced *P3 event-related potential* (ERP, for a review see Polich, 2007) that has also been shown to be sensitive to changes in vigilance (Koelega, Verbaten, van Leeuwen, Kenemans, Kemner, & Sjouw, 1992; Schmidt et al., 2009) and can be interpreted as a measure of the processing depth of the auditory stimulus.

1.4. Hypothesis

We hypothesized that (1) the drivers' vigilance state should decrease with driving distance (main-effect of distance); (2) the verbal assessment of drivers' state by the investigator should improve drivers' vigilance state (main-effect of communication).

2. Methods

2.1 Participants

Twenty-six right-handed participants (20 male, 6 female; age: M = 26.6, range: 21-40) with extensive driving experience (mean annual driving distance of approximately 20,500 km, i.e. 12,800 miles per year) were recruited on a voluntary basis for an "in-car EEG driving study". Participants were screened for a variety of exclusion criteria (handedness, auditory and visual disabilities, various illnesses), instructed to sleep regularly the night before the experiment, and to refrain from consuming caffeine in the morning on the day of the experiment. For their participation they received a monetary compensation of € 100.

The size of the sample available for data analysis was reduced by three subjects who aborted the drive due to heavy fatigue after less than 2:30 h driving duration which would have resulted in too few data points for a reliable analysis. As a result, 23 (18 male, 5 female; age: M = 26.7, range: 21-40) data sets containing all measures were available for statistical analysis.

2.2 Materials and Procedure

Participants arrived at 9:00 am and signed an informed consent form. While the physiological recording equipment was applied, the participants completed the German versions of the morningness-eveningness-questionnaire (D-MEQ: Griefhahn, Künemund, Bröde, & Mehnert, 2001) and the Edinburgh handedness inventory (Oldfield, 1971). Prior to the test-drive in real traffic the participants completed a monotonous simulated driving session of approximately 1.5 h duration starting at 10:00 am in a low-level fixed base simulator (Lane Change Task: Mattes, 2003). The results of the simulator study will be published elsewhere and are not subject of the present paper. To control for possible circadian (Folkard, 1997; Lenné, Triggs, & Redman, 1997) and nutritional effects (Smith & Miles, 1986) all participants had lunch at 11:30 am before a 30-minutes in-car EEG-baseline was recorded containing typical body movements (i.e. look over shoulder) and resting-EEG to gain experience on typical artifacts. Following the German legal maximum uninterrupted driving duration for commercial drivers (4.5 h) the length of the drive was set to approximately four hours and fifteen minutes. To ensure as much monotony as possible, the participants drove on a low-traffic highway off rush hour times (A 81 Autobahn, south of Stuttgart, Germany). To minimize the systematic influence of route-specific factors, we decided to vary the exits where the participants were to change direction during the drive. This resulted in two different experimental courses of 475 km (about 297 miles; tuning points at 118, 190, 261 and 342 km) and 481 km (301 miles; tuning points at 176, 248, 303 and 358 km).

Participants started at 12:45 pm and returned, on average, after 4:08 h of driving, except for cases in which the experiment was terminated earlier by the participants or the investigator. The final sample included 19 subjects that completed the whole drive and four subjects that dropped out earlier but still supplied enough data for a reliable analysis (average break-off after 3:04 h of driving).

Participants had to be sufficiently rested upon arrival. They were informed that they could stop driving at any time without any monetary or other penalties. An investigator accompanied the participant in the car, continuously monitoring the driver and ready to intervene whenever necessary. The test car was a Mercedes Benz S-Class (W221). The participants' task was to drive at a speed not exceeding 130 km/h (approximately 80 mph; recommended maximum speed on German highways) and to comply with the traffic rules at all times. They were instructed to use automatic shift and to refrain from turning on the radio or using other in-car devices. Further, participants were asked not to talk to the investigator outside the communication episodes except for the case of expressing the wish to abort the driving. Participants were also instructed to avoid unnecessary movements in order to reduce artifacts in the EEG recording. The participants knew that they would be prompted for their subjective state frequently throughout the drive. They were not informed about when and how often they were to be asked by the investigator nor did they know about the researchers' particular interest in the effect of the communication episodes.

2.3 Communication Episodes and Subjective Measures

In a prior research setting (Schmidt et al., 2009) the retrospective vigilance assessment occurred every 20 minutes. Despite this drivers experienced high levels of monotony and showed a pronounced vigilance decrement over time. We therefore decided to assess vigilance every 20 minutes. As subjective indicators we used the KSS (Åkerstedt & Gillberg, 1990) as a well-established single-item indicator of sleepiness and three similarly constructed items assessing inattention to the driving task, inattention to the auditory tones, and experienced monotony. The investigator verbally prompted the driver to judge his or her state concerning the four items on a nine point Likert-scale with regard to the previous 20 minutes of driving time. The typical sequence of the dialogue containing the four subjective dimensions is displayed in Table 1.

Table 1: Typical communication sequence

Investigator	Participant
"The next state assessment is up."	
	"O.K."
"Concerning the time period since the last prompting, how would you describe your predominant state on a scale from 1 -extremely alert- to 9 -extremely sleepy, fighting sleep-?"	
	"5"
"…how attentively have you been driving on a scale from 1 -extremely attentively- to 9 -extremely inattentively-?"	
	"4"
"…how attentively have you been reacting to the tones on a scale from 1 -extremely attentively- to 9 -extremely inattentively-?"	
	"4"
"…and how did you perceive the drive on a scale from 1 -extremely varied- to 9 -extremely monotonous-?"	
	"6"

2.4. Sections of Analysis

In order to define a baseline for each participant that should reflect alert driving, we analyzed the first twenty minutes, that is, the section before the first communication. This baseline is descriptively reported in the figures for all measures and in case of the alpha spindle rate was used to derive an individual index. Of all communication episodes only those that were not corrupted by any turning point entered the analysis because performing a turn might have had an activating influence. The remaining episodes were divided by half into early and late communications. To keep the amount of data entering the analysis for the early and late conditions constant within each participant, in case of an odd number of episodes the middle episode did not enter the final analysis. In order to define a stable state prior to the communications for all measures an interval from t = -5:30 min to t = -0:30 min was evaluated relative to the start of the communication episode. Although very unlikely, we left a gap of 30 s prior to the communication episode to rule out any possible anticipation of the communication by the participants that could have affected their state prior to the communication. For the continuous measures the communication interval per se was analyzed. Finally, an interval of five minutes directly following the communication was analyzed. For the continuous measures this interval was further divided into five sections of one-minute length each. For every dependent measure and every interval the mean over all available episodes was calculated.

2.5 Performance Measures

The participants were instructed that the auditory oddball reaction time task was only to be completed if they felt that it was safe to do so in a given driving situation. Infrequent (oddball-) tones (500 Hz, 20% probability) were presented in a random sequence mixed with frequent tones of a lower pitch (400 Hz, 80%). The inter-stimulus interval varied randomly between four and six seconds. The constantly low pitch of the stimuli reduced their alerting potential to a minimum. In a 2-AFC paradigm the participants had to respond to the tones by pressing buttons fitted to their right or their left thumbs. They were informed about their initial assignment of tone and response hand right before the start of the experiment (lower-pitched tone: left thumb, higher-pitched tone: right hand; or vice versa). This assignment was reversed half way through the drive for each participant. The initial assignment was selected randomly with the goal that, at the end of the experiment, the two types of assignment would have occurred equally frequently. Responses faster than 200 ms and slower than 4000 ms were not analyzed. For every analysis epoch, and separately for each participant, the mean of all reaction times above the 80th-percentile was calculated as a measure of the participants' slow reactions, and the mean of all reaction times below the 20th-percentile as participants' fast reactions.

2.6 Physiological Measures

Electroencephalogram (EEG), electrooculogram (EOG), and electrocardiogram (ECG) were recorded from 64 electrodes (500 Hz sampling rate, low cut-off: 0.016 Hz; high cut-off: 250 Hz) using BrainAmp recording hard- and software (Brainproducts GmbH, Munich, Germany). Data were down-sampled to 128 Hz and low-pass filtered at 48 Hz. Artefactual channels were excluded from further analysis. In order to minimize the influence of muscle activity, eye blinks and technical noise, we de-noised the data using the extended infomax ICA algorithm (Lee, Girolami, & Sejnowski, 1999), available in the EEGLab toolbox (Delorme & Makeig, 2004). We used an automated algorithm to extract monochromatic bursts of activity within the alpha band (7 - 13 Hz), which we call alpha spindles, and determined the amplitude, frequency, and duration of these activity bursts. In this paper we focus on the alpha spindle rate, defined as the number of alpha spindles per minute within each analysis epoch. Since the alpha rhythm is most prominent over parieto-occipital sites, we analyzed the signal of electrode Pz. To account for inter-individual differences an alpha spindle index was derived by dividing each section by the reference value of the first twenty minutes of driving.

To extract the P3 amplitudes the pre-processed EEG signal of electrode CPz was averaged time-locked to the presentation of the oddball stimulus. A baseline-correction (relative to -100 to 0 ms pre-stimulus time window) was applied. The P3 amplitude then was defined as the maximum value of the signal in a time window from 300 to 600 ms post-stimulus minus the average value of the baseline-epoch. The procedure was in accordance with the guidelines issued by Duncan, Barry, Connolly, Fischer, Michie, Näätänen et al. (2009).

In order to assess the duration of each eye blink in the vertical EOG, we identified its closing and reopening times in accordance with thresholds defined by Caffier et al. (2003) using an algorithm implemented in Matlab. Only blinks with durations between 50 and 500 ms and amplitudes larger than 100 µV entered the analysis. The mean duration over all blinks falling into the respective interval was calculated and entered into the statistical analysis.

R-peaks were identified from the ECG using an automated algorithm in Matlab. The average heart rate (beats per minute) was calculated for every interval.

2.7 Experimental Design

The independent variables in our experimental design were the occurrence of the communication episodes during the drive (driving segment: early vs. late) and the time relative to the communication (communication: two sections for the event-related measures [-5:30 to -0:30 and +0:00 to +5:00 min]; seven sections for the continuous measures [-5:30 to -0:30 min, during communication, and 5 sections of 1 minute length each following the communication]). Dependent variables were alpha spindle rate, heart rate, blink duration, slow and fast reactions, and P3 amplitude.

An a priori statistical power analysis using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) showed that in order to detect effects of $f = .50$ of the communication within-subject variable assuming a population correlation among the levels of this variable of $\rho = .50$ (estimated from pilot data) and given $\alpha = \beta = .05$, between $N = 14$ and $N = 16$ participants were needed (depending on the number of levels of the communication variable; see previous paragraph). A sensitivity power analysis showed that given a final sample of $N = 23$, effects of size $f = .30$ to $f = .39$ could be detected under otherwise identical conditions. A multivariate approach (MANOVA) was used for all within-subject comparisons to identify the effects of the driving segment and communication factors for each dependent measure. All multivariate test criteria correspond to the same (exact) F-statistic, which is reported. The level of α was set to .05 for all analysis. Partial η^2 is reported as a measure of relative effect size. Statistically significant results of the continuous measures were subjected to post-hoc contrast analysis using the PRE-section as the fixed reference. In these cases the level of α was Bonferroni-

Holm corrected (Holm, 1979) so as to avoid alpha error accumulation. For clarity of presentation only significant contrasts are reported.

3. Results

It was confirmed that all participants were right-handed (all handedness-indices > 0). The D-MEQ results showed that there were no participants falling into any of the extreme chronotype groups (moderate evening-type: 7; neutral type: 13; moderate morning-type: 3). On average about eight communication episodes entered the analysis for each participant (range: 4 to 12). The mean starting time of the early episodes was two hours before the mean starting time of the late ones. Comparison of early and late communication episodes revealed that the late episodes were significantly shorter (62 s (early) vs. 50 s (late); paired sample t-test: $t(22) = 5.19$; $p < .001$; $\eta^2 = .550$). For the subjective measures participants reported significantly higher judgments of all measures of vigilance for late as compared to early communication episodes (Figure 1; sleepiness: $t(22) = 4.20$; $p < .001$; $\eta^2 = .445$; inattention to driving: $t(22) = 4.37$; $p < .001$; $\eta^2 = .465$; auditory inattention: $t(22) = 3.78$; $p = .001$; $\eta = .394$; monotony: $t(22) = 3.86$; $p = .001$; $\eta^2 = .404$).

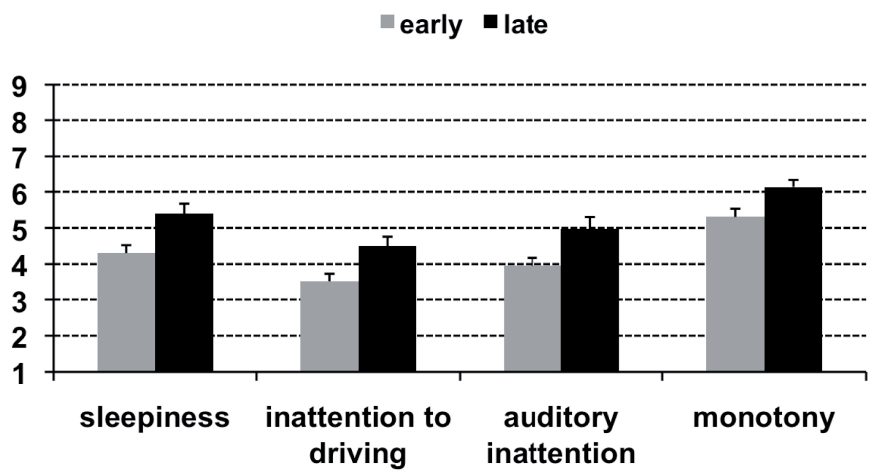

Figure 1: Subjective measures for early vs. late communication episodes.

3.1. Continuous Physiological Measures

The results of the MANOVAs testing the effects of driving segment and communication on the continuous physiological measures are summarized in Table 2. Both variables exerted a significant influence on alpha spindle rate and heart rate. Alpha spindle rate (Figure 2) was significantly higher around late as compared to early episodes. Relative to the pre-communication baseline, a significant reduction of alpha-spindle rate was observed during the communication episodes ($F(1,22) = 24.59$; $p < .001$; $\eta^2 = .528$). The reduction persisted for the subsequent minute ($F(1,22) = 18.24$; $p < .001$; $\eta^2 = .453$). The effect of driving segment on blink duration (Figure 3) just fell short of the preset criterion of statistical significance ($F(1,22) = 4.07$; $p = .056$). The average blink duration was significantly reduced during the communication ($F(1,22) = 17.99$; $p < .001$; $\eta^2 = .450$) relative to the pre-communication baseline. This effect persisted for two post-communication minutes (first minute: $F(1,22) = 17.46$; $p < .001$; $\eta^2 = .442$; second minute: $F(1,22) = 12.85$; $p = .002$; $\eta^2 = .369$). Heart rate (Figure 4) was significantly reduced around late as compared to early episodes. Relative to the pre-communication baseline, heart rate increased significantly during the communication ($F(1,22) = 44.86$; $p < .001$; $\eta^2 = .671$), but returned to the pre-communication level during the first post-communication minute. In the second post-communication minute, heart rate fell below the pre-communication level ($F(1,22) = 8.88$; $p = .007$; $\eta^2 = .287$).

Table 2: Statistical results for continuous physiological measures

	Driving Segment			Communication			Driv. Segm. x Communication		
	$F(1,22)$	p	η^2	$F(6,17)$	p	η^2	$F(6,17)$	p	η^2
Alpha Spindle Rate	15.84	<.001	.419	7.07	<.001	.714	1.02	.447	n.s.
Blink Duration	4.07	.056	n.s.	6.14	.001	.684	.71	.684	n.s.
Heart Rate	20.80	<.001	.486	10.68	<.001	.790	1.25	.329	n.s.

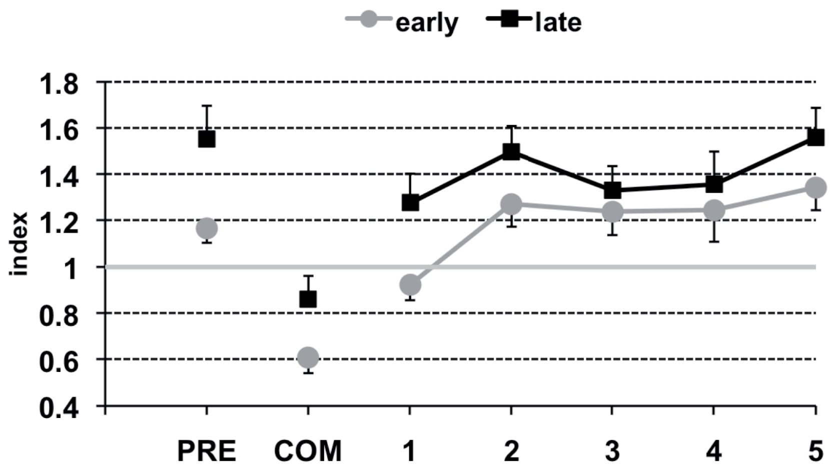

Figure 2: Alpha spindle rate prior to (PRE), during (COM), and in the minutes following the communication episode (1 to 5) for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first twenty minutes of driving.

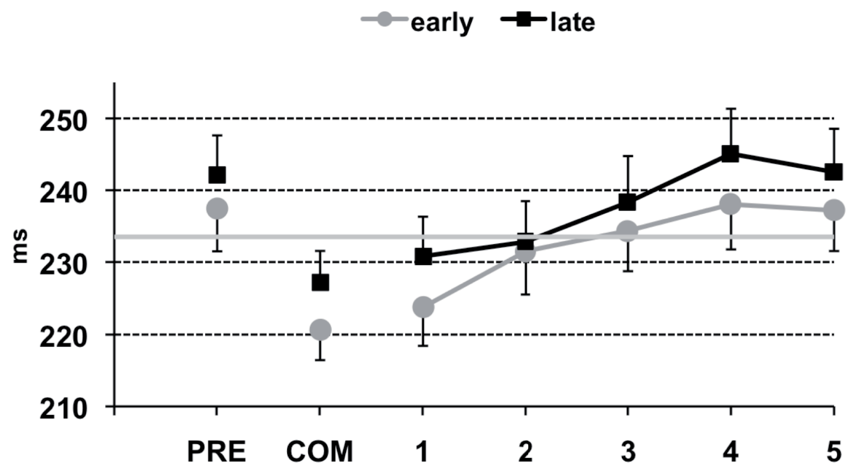

Figure 3: Blink duration prior to (PRE), during (COM), and in the minutes following the communication episode (1 to 5) for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first twenty minutes of driving.

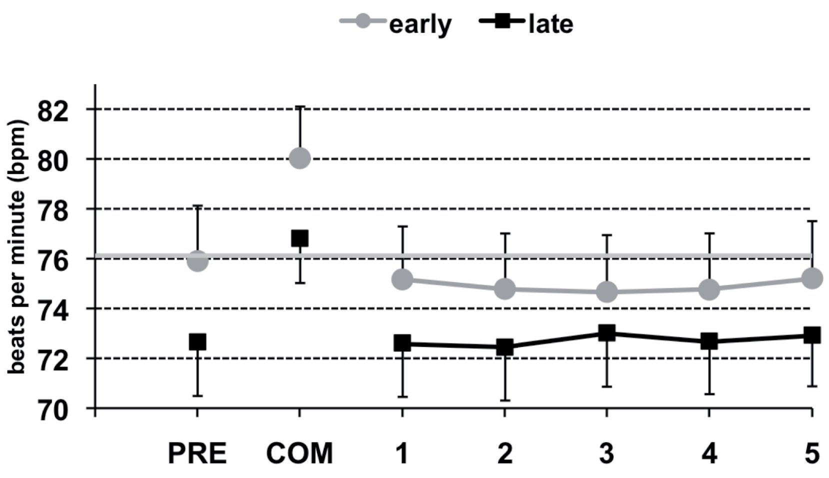

Figure 4: Heart rate prior to (PRE), during (COM), and in the minutes following the communication episode (1 to 5) for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first twenty minutes of driving.

3.2. Discrete Event-Related Measures

The results of the MANOVAs testing the effects of driving segment and communication on the discrete stimulus based measures are summarized in Table 3. As expected, we observed neither an effect of driving segment nor an interaction of driving segment and communication on the mean of the fast reaction times (Figure 5a). There was, however, a significant effect of the communication variable reflected in a decrease in the mean of the fast reaction times following the communication. In contrast, for the mean of the slow reaction times (Figure 5b) we observed significant main effects for both driving segment and communication with an increase in slow reaction times in later sections of the drive and a decrease in slow reaction times following a communication episode. There were no significant effects on the P3 amplitude (Figure 6).

Table 3: Statistical results for discrete event-related measures

	Driving Segment			Communication			Driv. Segm. x Communication		
	$F(1,22)$	p	η^2	$F(1,22)$	p	η^2	$F(1,22)$	p	η^2
Fast Reactions	3.03	.096	n.s.	10.16	<.001	.316	.39	.541	n.s.
Slow Reactions	15.27	<.001	.410	5.65	.027	.204	.53	.473	n.s.
P3 Amplitude	.00	.975	n.s.	.99	.330	n.s.	3.29	.084	n.s.

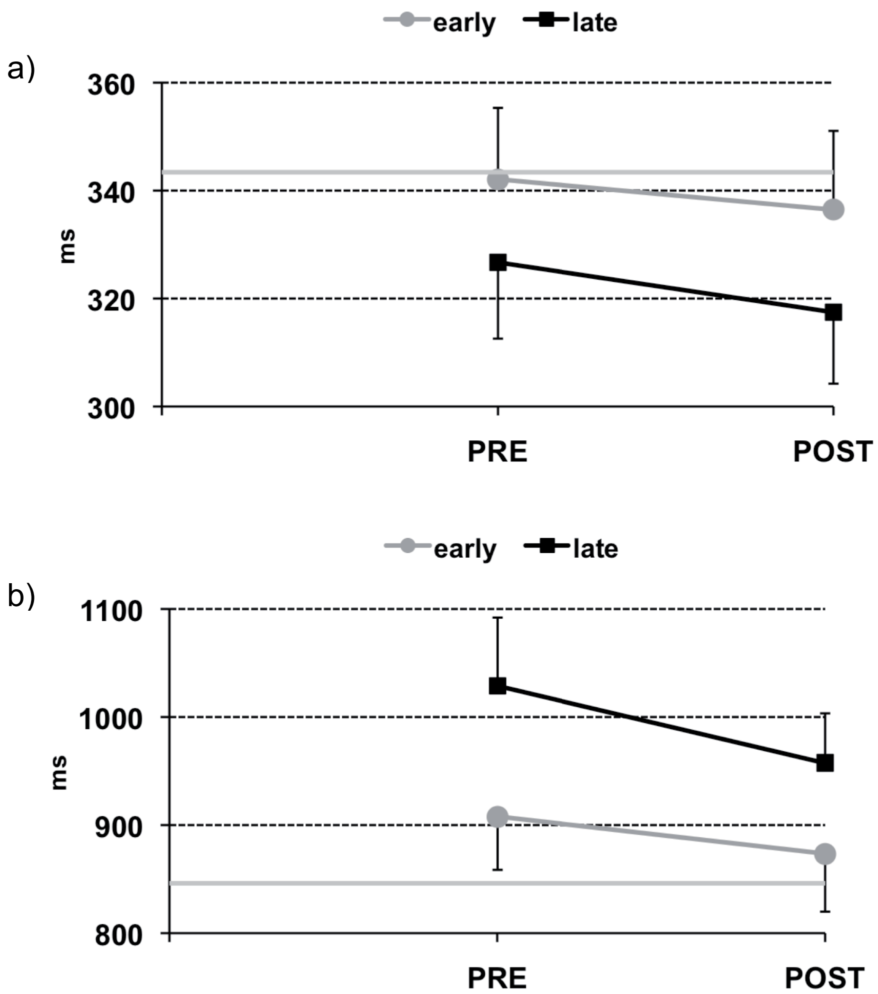

Figure 5: Means of fast (a) and slow (b) reaction times prior to (PRE) and following (POST) the communication episode for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal lines represent the baseline levels during the first twenty minutes of driving.

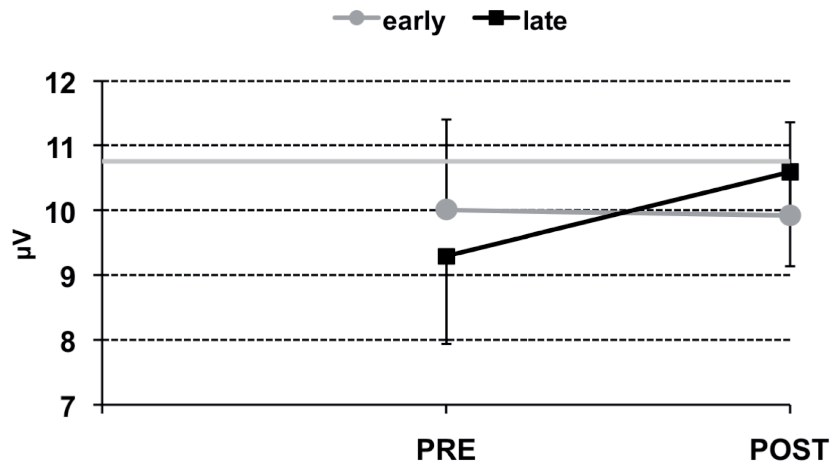

Figure 6: P3 amplitude prior to (PRE) and following (POST) the communication episode for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first twenty minutes of driving.

4. Discussion

The indices of subjective driver state as well as the objective vigilance measures - that is, alpha spindle rate, heart rate, and slow reaction times - all showed a significant effect of driving segment. This pattern replicates findings from an earlier study (Schmidt et al., 2009). Blink duration also showed a tendency towards an increased average duration with increasing distance driven. It may therefore be concluded that monotonous driving resulted in a vigilance decrement in the present study. We suspect the absence of this effect for the P3 amplitude (which was a sensitive measure of fatigue in the Schmidt et al. 2009 study) to be based on an insufficient number of trials entering the averaging procedure for each condition resulting in a poor signal-to-noise ratio.

During the communication episode all continuous measures indicated a clear improvement of vigilance in terms of a decrease in alpha spindle rate and blink duration and an increase in heart rate. This pattern suggests that a significant activation of the participants was induced by the communication. As mentioned in the introduction, it is not possible to infer, from the present data, the net effect of communication-induced driver activation on the positive side and distraction on the negative side. A reasonable assessment of this issue will likely require extended simulator studies in which critical variables such as brake reaction times can be studied in sufficient detail.

Concerning the duration of the effect elicited by the communication *following* the interaction, the continuous physiological measures lead to consistent results. Both alpha spindle rate and blink duration indicated an activating effect persisting after the end of the communication episode which, however, was limited to just one minute (alpha spindle rate) or two minutes (blink duration). Heart rate, in contrast, did not show any persisting effect and even moved below the pre-communication level after the communication. Thus, we conclude that the positive effects of verbal assessments on drivers' vigilance states may last beyond the communication episodes, but they do so only for a very short period of time.

Given this, we further conclude that an assessment interval of five minutes - the lower boundary recommended by Kecklund et al. (2006) - seems very reasonable if the goal is not to affect driving-induced fatigue. Strictly speaking, however, we cannot draw firm conclusions about possible cumulative effects of the repeated assessment (Kaida et al., 2007). This would have required running a control condition without any verbal assessments. Unfortunately due to its event-related nature the reaction time data did not allow us to reliably estimate the duration of the effect at a sufficient temporal resolution. Still, it is clear from the data that following the communication the means of slow as well as fast reactions decreased significantly. The effect for the fast reactions was rather surprising, because prior research had shown fast reactions to be insensitive to changes in vigilance (e.g., Schmidt et al., 2009). It may be that the interaction with the investigator induces self-awareness and as such draws the driver's attention to the fact that there is a reaction time task to be accomplished while driving. As a result, greater priority may be given to the secondary task, as a result of which even the fast reactions become faster.

A methodological issue inherent in the present study lies in the fact that, trivially, the post-communication epochs followed the pre-communication epochs. Therefore driving segment and communication were to some extent confounded, which could have resulted in a small underestimation of the activating effect of the communication episodes because vigilance decreased as a function of the distance driven. However, we believe that the amount by which the short-term activation was reduced by the underlying long-term fatigue is negligible.
Due to the fact that the communication process comprises various stages of cognitive and motor actions, future research is needed in order to analyze in more detail the factors that cause the activating effect. Candidates are the social interaction with another human being, the self-awareness induced by the requirement to think about ones own state, or the mere effect of talking or listening. Also, the degree to which the present results generalize to an everyday conversation while driving remains unresolved. For instance, on the one hand our participants only answered fairly routine questions, which may not be too activating per se. On the other hand, answering these questions may have been particularly activating due to their generating a certain degree of self-awareness. In real-world situations, the topics of the conversation may be much more variable (and likely more interesting) while at the same time being less self-referential. Finally, it is probably fair to say that considering in-car communications only in terms of the risks they imply without considering their positive effects - especially under conditions of monotonous driving conditions - is not appropriate, although the positive effects seem to be rather limited.

5. Conclusion

Verbal assessment of driver state causes a direct improvement of driver vigilance state as indicated by alpha-spindle rate, blink duration, and heart rate. The post-conversation temporal extension of this positive effect seems to be rather limited, as is indicated by the fact that the physiological vigilance measures returned to their pre-communication levels after two minutes at the latest. From these findings we conclude (a) that verbal assessments of fatigue are feasible even at a rate of one every five minutes, and (b) that there are clear positive effects of a verbal communication as a countermeasure for fatigue during the communication per se, but these positive effects can no longer be measured two minutes after the communication.

Acknowledgements

This research was supported by BMBF (German Federal Ministry of Education and Research) Grant 16SV2233. We thank Prof. Gabriel Curio, Dr. Ruth Schubert and Dr. Martin Fritzsche for valuable discussions on the experimental design; Andreas Pröttel and Sven Willmann for support with data recording and analysis; Michael Sachse and Ying Lee for their help with data collection.

References

Åkerstedt, T., & Gillberg M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 52, 29-37.

Anund, A., Kecklund, G., Peters, B., & Åkerstedt, T. (2008). Driver sleepiness and individual differences in preferences for countermeasures. Journal of Sleep Research, 17, 16-22.

Brookhuis, K., & De Waard, D. (1993). The use of psychophysiology to assess driver status. Ergonomics, 36 (9), 1099-1110.

Caffier, P.P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. European Journal of Applied Physiology, 89, 319-325.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. Accident Analysis and Prevention, 40, 1282-1293.

Chan, M., & Atchley, P. (2009). Effects of cell phone conversations on driver performance while driving under highway monotony. Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, MT, 140-146.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134, 9-21.

Drews, F.A., Pasupathi, M., & Strayer, D.L. (2008). Passenger and cell phone conversations in simulated driving. Journal of Experimental Psychology: Applied, 14(4), 392-400.

Drory, A. (1985). Effects of rest and secondary task on simulated truck-driving task performance. Human Factors, 27 (2), 201-207.

Duncan, C., Barry, R., Connolly, J., Fischer, C., Michie, P., Näätänen, R., Polich, J., Reinvang, I., & Van Petten, C. (2009). Event-related potentials in clinical research:

Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clinical Neurophysiology, 120 (11), 1883-1908.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.

Folkard, S. (1997). Black times: temporal determinants of transport safety. Accident Analysis and Prevention, 29 (4), 417-430.

Gershon, P., Ronen, A., Oron-Gilad, T., & Shinar, D. (2009). The effects of an interactive cognitive task (ICT) in suppressing fatigue symptoms in driving. Transportation Research Part F, 12, 21-28.

Graw, P., Kräuchi, K., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2004). Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. Physiology & Behavior, 80, 695-701.

Griefahn, B., Künemund, C., Bröde, P., & Mehnert, P. (2001). Zur Validität der deutschen Übersetzung des Morningness-Eveningness-Questionnaires von Horne und Östberg. Somnologie, 5, 71-80.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W.C. (1973). Quantification of sleepiness: a new approach. Psychophysiology, 10, 431-436.

Holm, S. (1979). A simple sequentially rejective multiple teste procedure. Scandinavian Journal of Statistics, 6, 65-70.

Horne, J., & Reyner, L. (1999). Vehicle accidents related to sleep: a review. Occupational and Environmental Medicine, 56, 289-294.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. Human Factors, 48, 196-205.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. Journal of Sleep Research, 15, 47-53.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical Neurophysiology, 117, 1574-1581.

Kaida, K., Åkerstedt, T., Kecklund, G., Nilsson, J.P., & Axelsson, J. (2007). The effects of asking for verbal ratings of sleepiness on sleepiness and its masking effects on performance. Clinical Neurophysiology, 118, 1324-1331.

Kecklund, G., & Åkerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. Ergonomics, 36, 1007-17.

Kecklund, G., Åkerstedt, T., Sandberg, D., Wahde, M., Dukic, T., Anund, A., & Hjälmdahl, M. (2006). State of the art review of driver sleepiness. DROWSI Project Report, Deliverable 1.1. Accessed 03.01.2010: www.vti.se/11934.epibrw

Koelega, H.S., Verbaten, M.N., van Leeuwen, T.H., Kenemans, J.L., Kemner, C., & Sjouw, W. (1992). Time effects on event-related brain potentials and vigilance performance. Biological Psychology, 34, 59-86.

Laurell, H., & Lisper, H.O. (1978). A validation of subsidiary reaction time against detection of roadside obstacles during prolonged driving. Ergonomics, 21, 81-88.

Lee, T.W., Girolami, M., & Sejnowski, T.J. (1999). Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources. Neural Computation, 11, 417-41.

Lenné, M.G., Triggs, T.J., & Redman, J.R. (1997). Time of day variations in driving performance. Accident Analysis and Prevention, 29 (4), 431-437.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In: H. Strasser, K. Kluth, H. Rausch and H. Bubb (Eds.), Quality of Work and Products in Enterprises of the Future, pp. 57-60, Ergonomia, Stuttgart.

May, J.F., & Baldwin, C.L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasures technologies. Transportation Research Part F, 12, 218-224.

Maycock, G. (1997). Sleepiness and driving: the experience of U.K. car drivers. Accident Analysis and Prevention, 29, 453-462.

Nordbakke, S., & Sagberg, F. (2007). Sleepy at the wheel: Knowledge, symptoms and behaviour among car drivers. Transportation Research Part F, 10, 1-10.

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

O'Hanlon J.F., & Kelly G.R. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In: R. Mackie (Ed.), Vigilance, pp. 87-109, Plenum Press, New York.

Oldfield, R.C. (1971). The assessment and analysis of handedness. The Edinburgh inventory. Neuropsychologia, 9, 97-113.

Oron-Gilad, T., Ronen, A., & Shinar, D. (2008). Alertness maintaining tasks while driving. Accident Analysis and Prevention, 40, 851-860.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P.D., Chouvarda, I., Bekiaris, E., & Maglaveras, N. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. Clinical Neurophysiology, 118, 1906-1922.

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Åkerstedt, T., Charles, A., & Bloulac, B. (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. Sleep, 28 (12), 1511-1516.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. Clinical Neurophysiology, 118 (10), 2128-2148.

Reyner, L.A., & Horne, J.A. (1997). Suppression of sleepiness in drivers: combination of caffeine with a short nap. Psychophysiology, 34, 721-725.

Reyner, L.A., & Horne, J.A. (1998). Evaluation of ‘‘in car’’ countermeasures to driver sleepiness: cold air and radio. Sleep, 21, 46-50.

Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics, 51 (7), 982-1010.

Schmidt, E.A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W.E. (2009). Drivers’ misjudgement of vigilance state during prolonged monotonous daytime driving. Accident Analysis and Prevention, 41, 1087-1093.

Smith, A.P., & Miles, C. (1986). The effects of lunch on cognitive vigilance tasks. Ergonomics 29 (10), 1251-1261.

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. Accident Analysis and Prevention, 35, 381-391.

Tietze, H., & Hargutt, V. (2001). Zweidimensionale Analyse zur Beurteilung des Verlaufs von Ermüdung. 43. Tagung experimentell arbeitender Psychologen, Regensburg, GER. Accessed 19.01.2010: http://www.psychologie.uni-wuerzburg.de/methoden/texte/2001_tietze_hargutt_Zweidimensionale_Analyse.pdf

Williams, H.L., Lubin, A., & Goodnow, J.J. (1959). Impaired performance with acute sleep loss. Psychological Monographs: General and Applied, 73 (14), 1-26.

Wright, N.A., Stone, B.M., Horberry, T.J., & Reed, N. (2007). A review of in-vehicle sleepiness detection devices. Published project Report 157. TRL Limited, UK, Crowthorne. Accessed 19.01.2010: http://www.masstransitmag.com/pdf/2007/mar/ABMT328Sleepy.pdf