

Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung

Berichte der
Bundesanstalt für Straßenwesen

Mensch und Sicherheit Heft M 203

The logo consists of the word "bast" in a bold, lowercase, green sans-serif font. The letters are slightly rounded and have a white outline, giving it a three-dimensional appearance. The logo is positioned in the bottom right corner of the page.

Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung

von

Sebastian Poschadel
Michael Falkenstein

Leibniz-Institut für Arbeitsforschung
an der Technischen Universität Dortmund

Preethy Pappachan
Eva Poll
Klaus Willmes von Hinckeldey

Universitätsklinikum Aachen
Neurologische Klinik,
Lehr- und Forschungsgebiet Neuropsychologie

**Berichte der
Bundesanstalt für Straßenwesen**

Mensch und Sicherheit Heft M 203

bast

Die Bundesanstalt für Straßenwesen veröffentlicht ihre Arbeits- und Forschungsergebnisse in der Schriftenreihe **Berichte der Bundesanstalt für Straßenwesen**. Die Reihe besteht aus folgenden Unterreihen:

A - Allgemeines
B - Brücken- und Ingenieurbau
F - Fahrzeugtechnik
M - Mensch und Sicherheit
S - Straßenbau
V - Verkehrstechnik

Es wird darauf hingewiesen, dass die unter dem Namen der Verfasser veröffentlichten Berichte nicht in jedem Fall die Ansicht des Herausgebers wiedergeben.

Nachdruck und photomechanische Wiedergabe, auch auszugsweise, nur mit Genehmigung der Bundesanstalt für Straßenwesen, Stabsstelle Presse und Öffentlichkeitsarbeit.

Die Hefte der Schriftenreihe **Berichte der Bundesanstalt für Straßenwesen** können direkt beim Wirtschaftsverlag NW, Verlag für neue Wissenschaft GmbH, Bgm.-Smidt-Str. 74-76, D-27568 Bremerhaven, Telefon: (04 71) 9 45 44 - 0, bezogen werden.

Über die Forschungsergebnisse und ihre Veröffentlichungen wird in Kurzform im Informationsdienst **BAST-Info** berichtet. Dieser Dienst wird kostenlos abgegeben; Interessenten wenden sich bitte an die Bundesanstalt für Straßenwesen, Stabsstelle Presse und Öffentlichkeitsarbeit

Impressum

Bericht zum Forschungsprojekt 82.291/2005:
Testverfahren zur psychometrischen
Leistungsprüfung der Fahreignung

Projektbetreuung

Hardy Holte

Herausgeber

Bundesanstalt für Straßenwesen
Brüderstraße 53, D-51427 Bergisch Gladbach
Telefon: (0 22 04) 43 - 0
Telefax: (0 22 04) 43 - 674

Redaktion

Stabsstelle Presse und Öffentlichkeitsarbeit

Druck und Verlag

Wirtschaftsverlag NW
Verlag für neue Wissenschaft GmbH
Postfach 10 11 10, D-27511 Bremerhaven
Telefon: (04 71) 9 45 44 - 0
Telefax: (04 71) 9 45 44 77
Email: vertrieb@nw-verlag.de
Internet: www.nw-verlag.de

ISSN 0943-9315
ISBN 978-3-86509-971-6

Bergisch Gladbach, Oktober 2009

Kurzfassung – Abstract

Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung

Ziel des Forschungsprojektes ist es, die Überprüfung der Fahreignung nach Fahrerlaubnis-Verordnung (FeV), Anlage 5, auf eine operationalere Basis zu stellen, die in der FeV und den Begutachtungs-Leitlinien Kap. 2.5 genannten Begriffe inhaltlich zu schärfen und Vorschläge für einen „Testkanon“ bzw. für Weiter- und Neuentwicklungen von standardisierten Tests zur Begutachtung zu erarbeiten.

Bei der Analyse der derzeit eingesetzten Testverfahren zur Überprüfung der psychometrischen Leistungsfähigkeit auf Basis der von der FeV geforderten wissenschaftlichen Standardisierung und Validierung zeigt sich, dass erheblicher Optimierungsbedarf besteht, insbesondere bei der Normierung und der Dokumentation der Testgütekriterien. In Einzelfällen fehlen sogar essenzielle Angaben. Die Herleitung der zu messenden Konstrukte ist ebenfalls oft unzureichend, teilweise gar nicht dokumentiert. Die Ergebnisse der kritischen Durchsicht wurden auf einem Workshop mit Experten aus Wissenschaft und Praxis diskutiert. Als Ergebnis des Forschungsprojektes lässt sich kurzfristiger und langfristiger Optimierungsbedarf festhalten.

Kurzfristig sollten die eingesetzten Testverfahren anhand vorher festgelegter Kriterien überarbeitet und an entsprechend großen repräsentativen Stichproben geeicht werden. Die Testgütekriterien müssen nachprüfbar angegeben werden.

Langfristig sollten aus wissenschaftlicher Perspektive die in der FeV, Anlage 5, genannten psychologischen Konstrukte grundsätzlich überarbeitet werden. Auch sollten die in der Fahreignungsdiagnostik angewendeten Entscheidungskriterien (Prozenträge) durch absolute Kennwerte ersetzt werden, die auf dem Wissen über die Zusammenhänge zwischen Leistung und Fahreignung basieren. Ein unabhängiges wissenschaftliches Gremium sollte die Güte und Zulassung der Testverfahren regelmäßig überprüfen.

In einem Teilprojekt wurde außerdem der Frage nachgegangen, inwieweit unterschiedliche Hardware- und Softwarekonstellationen die Objektivität der Reaktionszeitmessungen beeinflussen können. Die Ergebnisse zeigen, dass die zuverlässige Ermittlung von Reaktionszeiten in erster Linie von der

grundsätzlichen Eichung eines Testverfahrens abhängt. Computergestützte Leistungstests sollten deshalb in vom Hersteller zugelassenen Computerumgebungen im Durchschnitt bei unterschiedlichen Konfigurationen keinen Messfehler ausweisen, um einen Nachteil für Führerscheinbewerber auszuschließen. Außerdem müssten maximal mögliche Abweichungen und Konfidenzintervalle für zugelassene Konfigurationen angegeben werden, sowie Normwerttabellen vollständig einsehbar sein.

Tests for testing the psychometric performance of driver aptitude

The goal of the research project is to make sure, that testing of the driver aptitude in accordance with the German Driver's Licence Ordinance ("Fahrerlaubnis-Verordnung", FeV), annex 5, is operationalized more specifically. Furthermore, one should sharpen the mentioned terms within the German Driver's Licence Ordinance and sharpen the mentioned terms within the "Guidelines for Expertise on Driver Aptitude" Chapter 2.5 („Begutachtungs-Leitlinien zur Kraftfahreignung“). Moreover, recommendations should be made about a precise pool of tests and, if necessary, recommendations for further development of these tests for testing the driver aptitude in a standardized manner.

In accordance with the requested scientific "standardisation" and "validation" of the German Driver's Licence Ordinance an analysis was carried out. The outcome clearly shows, that an extensive optimization of the presently used tests for testing the psychometric performance of driver aptitude is necessary. This is especially the case for the standardisation of the examined tests and the documentation of the test quality factors ("Testgütekriterien"). In a few cases even essential information is missing.

The scientific definition of the constructs, which should be measured, are frequently inadequate and partly not even documented. This outcome was discussed by science and practical experts, invited by the BAST.

The experts and researchers came to the conclusion that there is still a great need of short and long term optimization:

In the near future tests should be revised on the basis of defined scientific criteria. Following they should be standardized, based on large representative samples. Furthermore, the test quality factors ("Testgütekriterien") must be described verifiably for every single test.

In the long perspective, seen from a scientific point of view, the psychological constructs mentioned in the German Driver's Licence Ordinance, annex 5, should be fundamentally revised.

Moreover, the decision criteria (percent distribution), applied in the "Guidelines for Expertise on Driver Aptitude", should be replaced by absolute variables based on the knowledge of the real relation between driving performance and driver aptitude. An independent scientific authority (consisting of more than one expert) should certify the quality of the tests (based on the determined scientific criteria) and decide about the general approval in regular time-related intervals.

In a subproject the question was raised, to which extend different hard- and software constellations could influence the objectivity of "reaction time" measurements. The results show that the reliable detection of response time depends first of all on the exact calibration of the test system in total.

Therefore, computer supported performance tests must not show any errors in measurement on average (for every soft- and hardware configuration, which is allowed by the manufacturer), to avoid a disadvantage of driver licence applicants. In addition, maximal possible errors in measurement must be declared, as well as completed tables of norm values (confident intervals included).

Inhalt

1	Einleitung	7	4.4	Diskussion der Ergebnisse in einer moderierten Expertengruppe	18
2	Problemstellung	7	4.5	Empfehlungen zur Anwendung der Begutachtungs-Leitlinien unter methodischen und neuropsychologischen Aspekten/Zusammenstellung angemessener Testverfahren	18
3	Stand der empirischen Forschung zur Kriteriumsvalidität von Verfahren zur psychometrischen Leistungsprüfung	11	5	Befragung von akkreditierten Begutachtungsstellen und ärztlichen Gutachtern	18
3.1	Allgemeine Literatur/Studien zur Fahreignungsüberprüfung	11	5.1	Fragebogenentwicklung	18
3.1.1	Eingesetzte Testverfahren	11	5.2	Durchführung der schriftlichen Befragung	19
3.1.2	Verwendete Methoden	12	5.3	Rücklaufquote und Stichprobenbeschreibung	19
3.1.3	Ergebnisse	12	5.4	Auswertung	20
3.1.4	Fazit – Forschungsergebnisse	14	5.4.1	Von den akkreditierten Begutachtungsstellen eingesetzte Verfahren ...	20
3.2	Angaben der Testhersteller zur Kriteriumsvalidität in den Testmanualen	14	5.4.2	Von Ärzten eingesetzte psychometrische Begutachtungsverfahren ...	20
3.2.1	ART 2020	14	5.4.3	Offene Fragen	20
3.2.2	Expertensystem Verkehr	15	5.4.4	Zusammenfassung der Auswertung ...	20
3.2.3	TAP-M	15	6	Beurteilung psychometrischer Tests zur Prüfung der Fahreignung	21
3.2.4	Corporal-A	15	6.1	Entwicklung einer Kriterienliste zur Beurteilung einzelner Tests	21
3.2.5	d2	15	6.2	Anmerkungen zur Reliabilitäts-schätzung bei computergestützten Reaktionstests	22
3.2.6	Konzentrations-Leistungs-Test	15	6.3	Vorgehensweise bei der Testanalyse ...	22
3.2.7	Zusammenfassung	15	7	Ergebnis der Testanalysen	23
3.3	Vergleich der Ergebnisse der empirischen Forschung und der Angaben der Testhersteller	16	7.1	Beschaffung der Testmanualen	23
3.4	Exkurs: das Kriterium der Fahrprobe	16	7.2	Testsystem A	24
3.5	Zusammenfassung	17	7.3	Testsystem B	25
4	Methodisches Vorgehen	17	7.4	Testsystem C	26
4.1	Schriftliche Befragung von Gutachtern über die in der Praxis eingesetzten Tests	17	7.5	Testsystem D	27
4.2	Auswertung der Testmanualen der erfragten und recherchierten Tests hinsichtlich Standardisierung	17	7.6	Test E	28
4.3	Einteilung der Tests in drei Kategorien	18	7.7	Test F	29
			7.8	Interpretation der Anforderungsbereiche in der Praxis durch verschiedene Testautoren	29

7.9	Einteilung der Tests in drei Kategorien	30	10.3	Ausgangslage	40
7.10	Fazit der Testanalysen	30	10.4	Methodisches Vorgehen/Untersuchungsdesign	42
8	Expertentreffen: Protokoll des Expertengesprächs zu den Anforderungen der Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung	30	10.4.1	Hardware	43
8.1	Teilnehmerübersicht	30	10.4.2	Software	43
8.2	Begrüßung und Vorstellungsrunde	31	10.4.3	Forschungsdesignplan	43
8.3	Problemstellung	31	10.5	Versuchsaufbau	43
8.3.1	Vorgeschichte des Projekts	31	10.6	Versuchsdurchführung	45
8.3.2	Zielsetzung des Projekts	31	10.7	Ergebnisse	45
8.4	Präsentation der Projekt-Ergebnisse	32	10.7.1	Ergebnisse im Detail	45
8.5	Diskussion der Ergebnisse	32	10.7.2	Einfluss der Einzelkomponenten	46
8.6	Optimierungsbedarf	36	10.7.3	Wechselwirkungen der Komponenten untereinander	47
8.6.1	Kurzfristige Strategie	36	10.8	Ermittlung der Normwerttabellen zur Überprüfung der Bedeutsamkeit der festgestellten Mittelwertabweichungen	48
8.6.2	Langfristige Strategie	37	10.8.1	Versuchsbeschreibung zur Ermittlung der dem untersuchten Testverfahren zu Grunde liegenden Normwerttabellen	48
9	Die Anforderungsbereiche der FeV unter neueren wissenschaftlichen aufmerksamkeits-theoretischen Aspekten	37	10.8.2	Ergebnisse zur Ermittlung der dem Test zu Grunde liegenden Normwerttabellen	48
9.1	Aktuelle Aufmerksamkeitstheorien	37	10.9	Abschließende Beurteilung des Einflusses unterschiedlicher Computerkonfigurationen auf die Reaktionszeitmessung	50
9.2	Welche Aufmerksamkeitsaspekte sind aus empirischer Sicht wichtig für die Fahreignung?	38	10.9.1	Grundsätzliche Eichung des Systems	50
9.3	Zuordnung der kriteriumsvalidierten Testverfahren zu den Anforderungsbereichen der FeV	38	10.9.2	Einfluss unterschiedlicher Hardware- und Softwarekonstellationen auf die Messung	51
9.4	Vorschlag zur Schärfung der derzeit bestehenden Anforderungsbereiche	39	10.9.3	Angabe von Prozentrangwerten und Konfidenzintervallen für die Auswertung	51
9.5	Vorschlag einer Testbatterie zur Fahreignungsprüfung	39	11	Zusammenfassung und Ausblick	51
10	Überprüfung der Reaktionszeitmessung eines computergestützten psychometrischen Testverfahrens	39	12	Literatur	55
10.1	Einleitung	39	12.1	Ergänzende Literatur „psychometrische Leistungstests zur Fahreignung und Fahrprobe“	58
10.2	Problemstellung	39			

1 Einleitung

Die Fahrerlaubnis-Verordnung (FeV), Anlage 5, legt Rahmenbedingungen für die Begutachtung von Bewerbern um die Erteilung oder Verlängerung der Führerscheinklassen D, D1, DE, DE1 sowie einer Fahrerlaubnis zur Fahrgastbeförderung fest. Wörtlich heißt es in Anlage 5 der FeV:

„1. Bewerber um die Erteilung oder Verlängerung einer Fahrerlaubnis der Klassen C, C1, CE, C1E, D, D1, DE, D1E sowie der Fahrerlaubnis zur Fahrgastbeförderung müssen sich untersuchen lassen, ob Erkrankungen vorliegen, die die Eignung oder die bedingte Eignung ausschließen. Sie haben hierüber einen Nachweis gemäß dem Muster dieser Anlage vorzulegen.

2. Bewerber um die Erteilung oder Verlängerung einer Fahrerlaubnis der Klassen D, D1, DE, D1E sowie einer Fahrerlaubnis zur Fahrgastbeförderung müssen außerdem besondere Anforderungen hinsichtlich:

- a. Belastbarkeit,
- b. Orientierungsleistung,
- c. Konzentrationsleistung,
- d. Aufmerksamkeitsleistung,
- e. Reaktionsfähigkeit

erfüllen.

Die zur Untersuchung dieser Merkmale eingesetzten Verfahren müssen nach dem Stand der Wissenschaft standardisiert und unter Aspekten der Verkehrssicherheit validiert sein.“

Derzeit fehlen konkrete Bestimmungen, welche psychometrischen Verfahren – und in welchem Umfang – Bestandteil der Begutachtung nach § 11, Abs. 9 und § 48 Abs. 4 und 5 FeV sein sollen.

Auch der Begriff der „wissenschaftlichen Standardisierung“ und „Validierung unter Aspekten der Verkehrssicherheit“ wird in der FeV, Anlage 5, nicht weitergehend definiert. Festgelegt werden allein die Bereiche, die begutachtet werden sollen. Das sind „Belastbarkeit“, „Orientierungsleistung“, „Konzentrationsleistung“, „Aufmerksamkeitsleistung“ und „Reaktionsfähigkeit“. In den „Begutachtungs-Leitlinien zur Kraftfahrereignung“ werden für die Begutachtung von Bewerbern zwar Prozenträge für einzelne Untertests festgelegt, doch welche Tests zu berücksichtigen sind oder wie eine mögliche Kom-

pensierbarkeit bei Unterschreiten der festgelegten Mindestanforderungen genau überprüft werden soll, bleibt offen und wird der Entscheidung des jeweiligen Gutachters überlassen: Es muss lediglich gewährleistet sein, dass zu jedem der begutachteten Bereiche wissenschaftliche, nachprüfbare Verfahren eingesetzt werden, die „unter Aspekten der Verkehrssicherheit validiert sein“ müssen. Darüber hinaus muss die Begutachtung nach der FeV von einem Betriebsarzt, einem Arzt für Arbeitsmedizin oder einer amtlich anerkannten Begutachtungsstelle für Fahreignung durchgeführt werden (FeV, Anlage 5, Abs. 2.).

Ziel des Forschungsprojektes ist es, die Überprüfung der Fahreignung nach FeV, Anlage 5, auf eine operationalere Basis zu stellen, die in der FeV und den Begutachtungs-Leitlinien genannten Begriffe inhaltlich zu schärfen und Vorschläge für einen „Testkanon“ bzw. für Weiter- und Neuentwicklungen von standardisierten Tests zur Begutachtung zu erarbeiten.

Das Forschungsprojekt wurde gemeinsam vom Institut für Arbeitsphysiologie an der Universität Dortmund (IfADo), Projektgruppe „Altern und ZNS-Veränderungen“, und dem Universitätsklinikum Aachen, Neurologische Klinik, Lehr- und Forschungsgebiet Neuropsychologie durchgeführt.

2 Problemstellung

Die Anlage 5 der Fahrerlaubnisverordnung (FeV) nennt 5 Anforderungen an die psychische Leistungsfähigkeit: Es sind dies (a) Belastbarkeit, (b) Orientierungsleistung, (c) Konzentrationsleistung, (d) Aufmerksamkeitsleistung und (e) Reaktionsfähigkeit. Diese Anforderungsbereiche bezeichnen unterschiedliche kognitive Bereiche, sind jedoch nicht ausreichend theoretisch basiert und entsprechen nicht neueren kognitiven Modellen (GOLZ, HUCHLER, JÖRG & KÜST, 2004).

Übliche Aufmerksamkeitstaxonomien unterscheiden Aspekte der Intensität und Selektivität. Intensität wird hierbei unterteilt in Alertness (allgemeine Reaktionsbereitschaft und deren kurzfristige Steigerung nach einem Warnreiz), intraindividuelle Aufmerksamkeitsschwankungen sowie längerfristige Aufmerksamkeitszuwendung. Der Selektivitätsaspekt umfasst fokussierte (nicht-räumliche) Aufmerksamkeit, geteilte Aufmerksamkeit und die Ausrichtung der Aufmerksamkeit im Raum (van ZOME-

REN & BROUWER, 1994). Zusätzlich wird eine bereichsübergreifende Funktion, das Supervisory Attentional System (SHALLICE, 1988), angenommen. Diese beinhaltet die Unteraspekte Strategie und Flexibilität (ZIMMERMANN & LECLERCQ, 2002).

Es konnte in funktionell-bildgebenden Untersuchungen wiederholt gezeigt werden, dass diese verschiedenen Aufmerksamkeitsaspekte spezifische und unterscheidbare funktionell-neuroanatomische Repräsentationen haben (GITELMAN et al., 1999). So sind z. B. Intensitätsaspekte (Alertness, allgemeine Reaktionsbereitschaft) vor allem mit rechtshemisphärischen und Hirnstamm-Strukturen (FERNANDEZ-DUQUE & POSNER, 2001; STURM et al., 1999; STURM & WILLMES, 2001), Aspekte der Aufmerksamkeitsausrichtung mit (vor allem rechts-)parietalen Strukturen (CORBETTA, KINCADE, OLLINGER, MCAVOY & SHULMAN, 2000; GITELMAN, PARRISH, FRISTON & MESULAM, 2002) sowie selektive nicht-räumliche Aufmerksamkeitskomponenten mit dem anterioren Gyrus cinguli und dorsolateralen präfrontalen Arealen (MILHAM et al., 2002; POSNER & RAICHLE, 1994) assoziiert. Dabei zeigte sich, dass der anteriore Gyrus cinguli in erster Linie mit Inhibition oder Faszilitierung von Reaktionen, der dorsolaterale präfrontale Kortex hingegen mit allgemeiner Aufmerksamkeitskontrolle (d. h. Zuweisung von Aufmerksamkeitskapazität, Fähigkeit, die Aufmerksamkeit aufzuteilen, „Top-down-Kontrolle“) verknüpft ist (KONDO, OSAKA & OSAKA, 2004; MILHAM, BANICH, CLAUS & COHEN, 2003).

Durch diese physiologischen Grundlagen, bei denen bestimmte kognitive Funktionen bestimmten Hirnarealen zugeordnet werden können, lassen sich auch die verschiedenen Einzelaspekte der Kognition besser voneinander separieren.

Derzeit werden auf Basis dieser Befunde drei neuronale Aufmerksamkeitsnetzwerke postuliert: das Alertness-Netzwerk, das Orientierungsnetzwerk und das Netzwerk der exekutiven Aufmerksamkeit (CALLEJAS, LUPIANEZ & TUDELA, 2004; FAN, McCANDLISS, SOMMER, RAZ & POSNER, 2002). Klinische Studien konnten nachweisen, dass diese Aufmerksamkeitskomponenten bzw. -netzwerke nach einer Hirnschädigung selektiv beeinträchtigt sein können (FIMM et al., 2001; HILDEBRANDT, GIESSELMANN & SACHSENHEIMER, 1999; KARNATH, FERBER & HIMMELBACH, 2001).

Diese drei Netzwerke stellen jedoch ebenfalls lediglich eine vereinfachte Konzeption der Aufmerksamkeit dar, die gemäß aktueller Befunde als vorläufig betrachtet werden muss. So betrachten NOBRE et al. (NOBRE, 2004; NOBRE, 2001) die Fokussierung als zentralen Aufmerksamkeitsaspekt, der eine Ausrichtung auf zeitliche, räumliche oder auch objektspezifische Merkmale beinhaltet. Hierbei muss zusätzlich nach der Verarbeitungsmodalität (visuell, auditiv, taktil) unterschieden werden, da gezeigt werden konnte, dass modalitätsspezifische Aufmerksamkeitsbeeinträchtigungen bei zerebraler Läsion auftreten können (OLSON, STARK & CHATTERJEE, 2003). Innerhalb jeder Modalität sind Aspekte der Aufmerksamkeitsintensität, der Fokussierung, der Inhibition sowie der räumlichen Aufmerksamkeitsausrichtung wirksam. Ein weiterer Bereich, der für unser Alltagsverhalten essenziell ist und vor allem in den letzten Jahren intensiver untersucht und als eigene Aufmerksamkeitskomponente betrachtet wird, ist die crossmodale Interaktion und Integration (SPENCE & DRIVER, 2004). Hiermit sind der Abgleich und die Integration von Informationen aus unterschiedlichen Sinneskanälen (z. B. visuell und auditiv) gemeint. So kann z. B. ein akustischer Reiz im Straßenverkehr zu einer plötzlichen Blickwendung (und möglicherweise zu einer Ablenkung vom Verkehrsgeschehen) führen.

Dem Netzwerk der exekutiven Aufmerksamkeit wird auch die Fähigkeit, die Aufmerksamkeit für verschiedene Aufgaben zu teilen, zugerechnet. Diese wird also als eine zentrale Fähigkeit bei der Handlungsregulation und Steuerung im Alltag und im Straßenverkehr betrachtet (BROUWER, 2002).

Weiterhin zeigen Querschnittsuntersuchungen zur Altersabhängigkeit verschiedener Aufmerksamkeitsaspekte ein lineares Ansteigen der Bearbeitungszeit und der intra- und individuellen Reaktionsvariabilität mit zunehmendem Alter, das besonders bei visuell-räumlichen sowie Prozessen, die andere Prozesse steuern (exekutive Aufmerksamkeitsfunktionen), ausgeprägt ist (ZIMMERMANN & FIMM, 2002a; ZIMMERMANN & FIMM, 2002b). Dies ließ sich sowohl in Verhaltensexperimenten (HASHER, STOLTZFUS, ZACKS & RYPMA, 1991; HOMMEL, LI & LI, 2004; TSANG & SHANER, 1998; UTTL & PILKENTON-TAYLOR, 2001) als auch mittels funktioneller Bildgebung (HEIN & SCHUBERT, 2004; MILHAM et al., 2002) bestätigen.

Die Anforderungsbereiche der FeV, Anlage 5, erweisen sich unter diesen Voraussetzungen als un-

scharf und teilweise redundant (z. B. Aufmerksamkeitsleistung, Konzentrationsleistung) und nicht operationalisiert.

Zwar beschreiben die Begutachtungs-Leitlinien mögliche Auswirkungen psychischer Leistungsmängel als Erläuterung der 5 o. g. Anforderungsbereiche, diese lassen sich jedoch nur eingeschränkt mit aktuellen Modellen der Aufmerksamkeit (und den zugrunde liegenden neuronalen Repräsentationen) in Einklang bringen und mit den 5 Anforderungsbereichen verknüpfen.

Tabelle 1 enthält den Versuch einer Zuordnung der 5 Anforderungsbereiche nach FeV zu den in den Begutachtungs-Leitlinien genannten Verhaltensauswirkungen und den aus den oben aufgeführten aktuellen Modellen abgeleiteten Aufmerksamkeitsfunktionen durch die Autoren. Vor allem die Anforderungsbereiche „Reaktionsfähigkeit“ und „Konzentrationsfähigkeit“ lassen sich mehreren Aufmerksamkeitskomponenten sowie Verhaltensaus-

wirkungen zuordnen. Zudem ist der Begriff „Aufmerksamkeit“ als eigentlicher Oberbegriff der verschiedenen Aufmerksamkeitskomponenten hier wenig aussagekräftig und möglicherweise eher in Bezug auf Aufmerksamkeitsverteilung bzw. exekutive Aufmerksamkeit gemeint.

Die im Rahmen der Fahreignungsuntersuchungen eingesetzten psychometrischen Testverfahren prüfen unterschiedliche der in Tabelle 1 genannten Aufmerksamkeitsaspekte, häufig mehrere gleichzeitig. Sie liefern in der Regel auch mehrere Normwerte, die unter Umständen unterschiedliche Anforderungen (z. B. Geschwindigkeit/Reaktionsfähigkeit vs. Sorgfaltsleistung) repräsentieren. Weiterhin sind die Testverfahren hinsichtlich ihrer Darbietungsform (Papier- und Bleistiftverfahren, apparative und computergestützte Testverfahren), ihrer Anforderungen an das visuelle Diskriminationsvermögen (durch die Größe und den Kontrast der verwendeten Stimuli bedingt), der geprüften sensorischen Modalität (visuell, akustisch, visuell-akus-

Anforderungsbereich (FeV)	Verhaltensauswirkung (Leitlinien, Kap. 2.5)	Aufmerksamkeitsfunktion
Visuelle Wahrnehmung; Reaktionsfähigkeit	1 Optische Informationen werden in ihrem Bedeutungsgehalt nicht ausreichend schnell und sicher wahrgenommen.	Alertness; allgemeine Reaktionsbereitschaft sowie darüber hinaus allgemeine sensorische Verarbeitung
Orientierung	2 Die Zielorientierung im jeweiligen optischen Umfeld, d. h. im Verkehrsraum, gelingt nicht oder nicht sicher oder nur mit einem so deutlich erhöhten Zeitaufwand, dass daraus in der konkreten Verkehrssituation eine Gefährdung entstehen würde.	Selektive, visuell-räumliche Aufmerksamkeit, d. h. Verschiebung des räumlichen Aufmerksamkeitsfokus
Konzentrationsfähigkeit	3 Die Konzentration ist zeitweilig oder dauernd gestört in der Weise, dass die jeweils anstehende Fahraufgabe aufgrund von Abgelenktsein oder Fehldeutungen verkannt oder fehlerhaft gelöst wird.	Selektive nicht-räumliche Aufmerksamkeit; Interferenzanfälligkeit; Ausblenden irrelevanter Informationen
Aufmerksamkeit	4 Die Aufmerksamkeitsverteilung ist unzulänglich, weil nur ein Teilbereich der für den Kraftfahrer bedeutsamen Informationen erfasst wird und/oder bei Situationswechsel, z. B. nach einer Phase der Monotonie, neue Informationen der Aufmerksamkeit entgehen.	Exekutive Aufmerksamkeit (geteilte Aufmerksamkeit; Umstellungsfähigkeit)
Belastbarkeit	5 Die Aufmerksamkeitsbelastung ist zu gering, weil es unter Stress oder nach länger andauernder Beanspruchung zu fehlerhaften Wahrnehmungen, Interpretationen oder Reaktionen kommt.	Aufmerksamkeitsintensität (längerfristige Aufmerksamkeitszuwendung)
Reaktionsfähigkeit	6 Notwendige motorische Reaktionen setzen zu spät ein und/oder werden stark verzögert ausgeführt.	Alertness (allgemeine Reaktionsbereitschaft)
Reaktionsfähigkeit; Konzentrationsfähigkeit	7 Reaktionen erfolgen unsicher, eventuell vorschnell und situationsunangemessen oder werden unpräzise, motorisch ungeschickt, „überschießend“ oder überhastet ausgeführt.	Selektive nicht-räumliche Aufmerksamkeit; Inhibition unerwünschter Reaktionen; Reaktionsselektion; motorische Präzision
Konzentrationsfähigkeit	8 Die psychischen Leistungen sind instabil in dem Sinne, dass die erforderliche Ausgewogenheit zwischen Schnelligkeit und Sorgfaltsleistung fehlt.	Selektive Aufmerksamkeit; speed-accuracy trade-off (Balance zwischen Geschwindigkeit und Sorgfalt)

Tab. 1: Mögliche Zuordnung der in den Begutachtungs-Leitlinien (BAST, 2000) genannten psychischen Verhaltensauswirkungen zu aktuellen Modellen der Aufmerksamkeit durch die Autoren

tisch), der motorischen Anforderungen (durchstreichen, Tastendruck, Fußpedal etc.) und der Komplexität der Testinstruktion sehr heterogen. Erschwerend kommt hinzu, dass sich die Tests hinsichtlich der verwendeten Normstichproben (Größe, Alter und Relevanz für die Fahreignungsprüfung) sowie im Ausmaß vorliegender Validierungsuntersuchungen in Bezug auf Fahrtauglichkeit unterscheiden. Zudem konnte gezeigt werden, dass häufig verwendete Testverfahren eine z. T. nur mäßige vorhersagende Validität gemessen an einem globalen Verhaltensrating nach einer Fahrprobe aufweisen (HANNEN, HARTJE & SKRECZEK, 1998).

Schon die kleine Auswahl an Testverfahren in Tabelle 2 und deren mögliche Einordnung bzgl. FeV-Anforderungsbereich machen deutlich, dass eine eindeutige Zuordnung der Tests zu den Anforderungsbereichen nur in Ansätzen möglich ist (für den Corporal-A konnte nach Meinung der Autoren wegen mangelnder theoretischer Einbettung im Testmanual keine zuverlässige Zuordnung erfolgen). Vielmehr vereinen einige Tests mehrere unterschiedliche Anforderungen (z. B. Wiener Determinationstest), andere wiederum weisen ein relativ spezifisches Aufmerksamkeitsprofil auf (z. B. Wiener Reaktionstest/Licht). Dies führt zu folgenden Überlegungen:

1. Die in der FeV genannten 5 Anforderungsbereiche sollten inhaltlich präzisiert werden im Hinblick auf darunter subsumierte, empirisch experimentell und funktionell-bildgebend nachgewiesene Aufmerksamkeitsfunktionen. Dies ermöglicht in der Folge eine präzisere Aussage da-

rüber, welcher Anforderungsbereich mit welchen Tests am besten erfasst werden kann.

2. Aufgrund der unterschiedlichen funktionalen Spezifität der psychometrischen Tests ergibt sich eine möglicherweise unterschiedliche Relevanz der Verfahren zur Vorhersage der Fahreignung. Diese hängt davon ab, welche Aufmerksamkeitskomponenten jeweils in die Testleistung eingehen und wie valide diese Komponenten unter Aspekten der Verkehrssicherheit sind. In diesem Zusammenhang plädiert BROUWER (BROUWER, 2002) u. a. für die Verwendung von Aufgaben zur geteilten Aufmerksamkeit bei der Fahreignungsprüfung.
3. Je breiter das Anforderungsprofil eines Tests, desto größer ist die Wahrscheinlichkeit, dass spezifische Defizite innerhalb des Verfahrens kompensiert werden können (z. B. ein sehr niedriger Prozentrang in Parameter 1 durch einen hohen Prozentrang in Parameter 2).

Diese Aspekte haben direkte Auswirkungen auf die Gewichtung und Interpretation der Ergebnisse einer Fahreignungsprüfung im Einzelfall.

Für das methodische Vorgehen im Forschungsprojekt bei der Untersuchung den bestehenden Testverfahren bedeutet dies:

1. Neben der inhaltlichen Analyse der zur Fahreignung eingesetzten Testverfahren zur Beurteilung der daran beteiligten Aufmerksamkeitskomponenten sind Analysen der Interkorrelationen dieser Verfahren (sofern publiziert) relevant, um

Anforderungsbereich	Aufmerksamkeitsfunktion	Computergestützte /apparative Tests	Papier- und Bleistifttests
Belastbarkeit	Längerfristige Aufmerksamkeit bei hoher Beanspruchung	Wiener Determinationsgerät, Wiener Testsystem – Daueraufmerksamkeit	
Orientierung	Visuell-räumliche Aufmerksamkeit ohne bzw. mit Blickbewegungen	Linienverfolgungstest, TAPK-Visuelles Scanning, ART-2020, TAVT 2, Wiener Determinationsgerät	Zahlenverbindungstest (ZVT)
Konzentration	Selektive, visuelle, nicht-räumliche Aufmerksamkeit	Wiener Testsysteme – Cognitrone, Wiener Reaktionstest (Ton bzw. Licht)	Aufmerksamkeits-Belastungstest d2 nach BRICKENKAMP, Frankfurter Aufmerksamkeitsinventar (FAIR), Farbe-Wort-Interferenztest
Aufmerksamkeit	Geteilte Aufmerksamkeit Umstellungsfähigkeit	TAVT, 2 TAPK-Geteilte Aufmerksamkeit, TAPK-Reaktionswechsel	
Reaktionsfähigkeit	Alertness (allgemeine Reaktionsbereitschaft), Sicherheit und Präzision motorischer Reaktionen	Wiener Reaktionsgerät (Ton bzw. Licht), Wiener Determinationstest	

Tab. 2: Mögliche Zuordnung ausgewählter psychometrischer Fahreignungstests zu den 5 Anforderungsbereichen der FeV sowie zu Aufmerksamkeitsfunktionen durch die Autoren

Hinweise zu deren inhaltlicher (Aufmerksamkeits-)Spezifität sowie möglicher Redundanz zu erhalten.

2. In einem weiteren Schritt ist die Analyse der Validität der Testverfahren unter Aspekten der Verkehrssicherheit – d. h. der Kriteriumsvalidität im Hinblick auf reales Fahrverhalten (z. B. Gesamturteil, Verhalten in spezifischen Verkehrssituationen) bzw. anderer objektiver Informationen (z. B. Unfallhäufigkeit) – von besonderer Relevanz. Dies ermöglicht schließlich die Aussage darüber, welche Aufmerksamkeitskomponenten in welchem Ausmaß zur Vorhersage der Fahreignung beitragen.
3. Dieses Vorgehen soll dann schließlich zu einer inhaltlichen Schärfung der Anforderungsbereiche und einer Präzisierung diagnostischer Kriterien führen.

3 Stand der empirischen Forschung zur Kriteriumsvalidität von Verfahren zur psychometrischen Leistungsprüfung

Aufbauend auf einer Sichtung der aktuellen Literatur wurde ein Überblick zum Thema „Kriteriumsvalidität von Testverfahren zur psychometrischen Leistungsprüfung“ erstellt. Die Ergebnisse der Recherche werden in diesem Kapitel dargestellt. Es wird ausdrücklich darauf hingewiesen, dass bei dieser Darstellung kein Anspruch auf Vollständigkeit erhoben werden kann. Es wird Bezug auf Veröffentlichungen genommen, die in den unten genannten Datenbanken gefunden wurden und zugänglich waren, sowie auf die in den Testmanualen der Hersteller genannten Informationen.

In einem Exkurs wird ferner darauf eingegangen, wie das Kriterium der Fahrprobe in aktuellen Studien definiert wird und mit welchen Herausforderungen die Festlegung der Bedingungen einer Fahrprobe einhergeht.

3.1 Allgemeine Literatur/Studien zur Fahreignungsüberprüfung

Zur Überblicksgewinnung über den aktuellen Stand der Forschung wurde eine Literaturrecherche in den Datenbanken PubMed, PsycINFO und PSYNDEX sowie im Internet über den Dienst „Google

Scholar“ durchgeführt. Eingeschlossen wurden alle Veröffentlichungen, die bis Mitte Juni 2007 in die Datenbanken eingegangen sind. Über diese Recherche konnten insgesamt 167 Artikel identifiziert werden, die für die neuropsychologische Bewertung der Fahreignung von Interesse sind. Eine Liste der Literatur findet sich im Kapitel 12.1 „Ergänzende Literatur“ am Ende des Berichtes. Von diesen Artikeln behandeln 19 die praktische Fahrprüfung als Gegenstand der Fahreignung, hiervon 9 aus Deutschland, Österreich und Belgien (AKINWUNTAN et al., 2006; AKINWUNTAN et al., 2002; BUKASA & PIRINGER, 2001; BUKASA, CHRIST, PONOCNY-SELIGER, SMUC & WENNINGER, 2003B; De RAEDT & PONJAERT-KRISTOFFERSEN, 2000; HANNEN et al., 1998; HELD, LAMBERTI & KUBITZKI, 1993; KROLL et al., 2003; SOMMER & HÄUSLER, 2006). Eine Studie (KARNER & BIEHL, 2001) verglich im Sinne einer Kreuzvalidierung Testsysteme untereinander, sie wird daher zusätzlich aufgeführt. Im Folgenden wird ein Überblick gegeben über die eingesetzten Testverfahren, statistischen Methoden und die Ergebnisse der Studien.

3.1.1 Eingesetzte Testverfahren

Bei den Testverfahren ist erwartungsgemäß eine deutliche Unterscheidung zwischen Studien aus dem deutschen Sprachraum und internationalen Studien zu treffen. In Deutschland und Österreich wurden vor allem die Verfahren Expertensystem Verkehr (sowie die Vorgänger Wiener Determinationsgerät und Wiener Reaktionsgerät), ART 90 und 2020, Tachistoskopischer Verkehrsauffassungstest (TAVT), Aufmerksamkeits-Belastungs-Test d2, Testbatterie zur Aufmerksamkeitsprüfung zur Überprüfung der Fahrtauglichkeit (TAP-M, zum Teil nur Untertests), aber auch Linien-Verfolgungs-Test (LVT) und Trailmaking-Test (TMT) in Untersuchungen berücksichtigt. In internationalen Studien wurden besonders der TMT, das UFOV (Useful Field Of View) und die Rey-Osterrieth-Figur eingesetzt, daneben aber auch unterschiedlichste Verfahren zu Merkfähigkeit, Intelligenz, Aufmerksamkeit und Demenzdiagnostik.

Die Auswahl der gewählten Testverfahren erfolgte in den meisten Studien nicht theoriegeleitet, sondern richtete sich vor allem nach dem gängigen Einsatz dieser Testverfahren in der Aufmerksamkeits- und auch Fahreignungsdiagnostik beziehungsweise bei Patientenstichproben (z. B. Schä-

del-Hirn-Trauma, Demenz, Parkinson) nach für diese Patientengruppen etablierten Verfahren.

3.1.2 Verwendete Methoden

Die Studien unterscheiden sich nicht nur in der Auswahl der verwendeten psychometrischen Testverfahren, sondern auch im Hinblick auf die eingesetzten statistischen Methoden zur Vorhersage der Fahreignung bzw. zur Analyse des Zusammenhangs zwischen Testleistung und realer Fahrleistung in der Fahrprobe. Häufig werden Korrelationen zwischen einzelnen Tests und der Fahrprobe und Gruppenvergleiche (ANOVA, t-Test, U-Test) eingesetzt. Zusätzlich kommen Regressionsanalysen (z. B. KROLL et al., 2003), Diskriminanzanalysen (HANNEN, HARTJE & SKRECZEK, 1998; SINGH, PENTLAND, HUNTER & PROVAN, 2007), Pfadanalysen (BUKASA, CHRIST, PONOCNY-SELIGER, SMUC & WENNINGER, 2003a) und künstliche neuronale Netze (SOMMER, ARENDASY, SCHUH-FRIED & LITZENBERGER, 2006) zur Anwendung, allerdings – bis auf zwei Ausnahmen – nur in Studien aus Deutschland und Österreich. Die im Einzelnen verwendeten Methoden werden in den folgenden Kapiteln angegeben.

Zu den beschriebenen Methoden muss kritisch angemerkt werden, dass in der Praxis nur solche Verfahren sinnvoll einzusetzen sind, die dem Anwender eine direkte Aussage über die psychometrische Leistungsfähigkeit bezogen auf die Fahreignung des Probanden ermöglichen. Der bloße statistische Zusammenhang eines Testverfahrens zur psychometrischen Leistungsprüfung der Fahreignung reicht zur Beurteilung in der Praxis nicht aus: Vielmehr muss die Beurteilung der Fahreignung über Cut-offs oder andere dichotome oder künstlich dichotomisierende Variablen ermöglicht werden. Studien, in denen beispielsweise mit Korrelationen gearbeitet wird, haben daher nur eine begrenzte Aussagekraft hinsichtlich dieses Unterscheidungsmerkmals.

Die Methode des Extremgruppenvergleichs ist ebenfalls problembehaftet, da hiermit – durch ein Ausklammern des mittleren Bereichs der Variablenausprägung – in der Regel die Bedeutung der untersuchten Zusammenhänge überschätzt wird (BORTZ & DÖRING, 2003).

3.1.3 Ergebnisse

Obwohl die Studien sogar bei gleichen psychometrischen Testverfahren im Ergebnis zu unterschied-

lichen Aussagen gelangen, können für den Raum Deutschland/Österreich/Belgien folgende Tests positive Zusammenhänge zwischen den Werten im psychometrischen Leistungstest mit in einer Fahrprobe erbrachten Leistungen verzeichnen: Expertensystem Verkehr und Vorgänger Wiener Determinationsgerät (WDG) und Wiener Reaktionsgerät (WRG) (HANNEN et al., 1998; HELD et al., 1993; KROLL et al., 2003; SOMMER et al., 2006), ART 2020 (BUKASA et al., 2003a; BUKASA et al., 2001), TAP-M (AKINWUNTAN et al., 2006; AKINWUNTAN et al., 2002; De RAEDT et al., 2000; HELD et al., 1993) und die Rey-Osterreith-Figur (AKINWUNTAN et al., 2006; AKINWUNTAN et al., 2002). Im Einzelnen ergab sich für diese Verfahren folgendes Bild:

Expertensystem Verkehr

In einer einzelnen Studie mit 222 Teilnehmern in Wien und Bad Tölz wurden ausgewählte Variablen der Testbatterie Expertensystem Verkehr mittels Fahrprobe und der Methode künstlicher neuronaler Netze untersucht (Variablen: Überblicksgewinnung aus dem TAVT, Allgemeine Intelligenz aus dem Adaptiven Matrizen-Test Form S2, die Variable Richtige aus dem Determinationstest Form S1, Motorische Zeit und Reaktionszeit des Reaktionstests Form S3, die Mittlere Zeit korrekter Zurückweisung im Cognitron Form S1 sowie das Gesichtsfeld und in der Version 'Plus' die Trackingabweichung aus der Peripheren Wahrnehmung), (SOMMER et al., 2006). Mit Hilfe der Testbatterie Standard werden 73,13 Prozent der in der standardisierten Fahrprobe positiv bewerteten Personen anhand ihrer Testergebnisse mit einer Wahrscheinlichkeit $\geq 0,70$ dieser Gruppe zugeordnet. 13,63 Prozent der in der standardisierten Fahrprobe negativ bewerteten Personen wird mit einer Wahrscheinlichkeit $\geq 0,70$ fälschlich ein hinreichend sicheres Fahrverhalten unterstellt. Im Falle der Testbatterie Plus werden 76,87 Prozent der tatsächlich positiv bewerteten Personen mit einer Wahrscheinlichkeit $\geq 0,70$ dieser Gruppe zugeordnet, während 10,20 Prozent der tatsächlich negativ Beurteilten fälschlicherweise mit einer Wahrscheinlichkeit $\geq 0,70$ dieser Gruppe zugeordnet werden (SOMMER et al., 2006).

Zu den Ergebnissen der Validierung der Vorgängerversionen „Wiener Determinationsgerät“ und „Wiener Reaktionsgerät“ sei auf die Literatur verwiesen (HANNEN et al., 1998; HELD et al., 1993; KROLL et al., 2003).

ART 2020

Das ART-2020-Testsystem wurde in umfangreichen Messungen mit 120 (BUKASA et al., 2001) beziehungsweise 123 Teilnehmern (BUKASA et al., 2003a) in Wien untersucht. In der Fahrprobe kam das System zur Analyse des Fahrverhaltens (SAF) zum Einsatz, neben dem Fahrlehrer war ein unabhängiger Bewerter bei der Fahrt anwesend. Das SAF (SMUC, CHRIST & GATSCHA, 2006) erfasst per Kameraaufzeichnungen Fahrer, Fahrstrecke und Pedalbereich des Kfz, es speichert Fahrgeschwindigkeit, Beschleunigungen und Kfz-Statussignale. Der Beobachter gibt während dieser Aufzeichnung über eine Tastatur Beobachtungsvariablen wie zum Beispiel Spurhalten, Abstandhalten etc. ein. Diese Bewertungen können dann der zu dem Zeitpunkt aufgezeichneten Fahrsituation zugeordnet werden. In der Studie von 2001 wurden Korrelationen der Fahrprobevariablen mit den Testergebnissen berechnet sowie Extremgruppenvergleiche durchgeführt. In der Studie von 2003 wurden zuerst Faktorenanalysen berechnet (zum einen über die Fahrleistungsvariablen, zum anderen über die Testvariablen), hiernach wurde eine Pfadanalyse zur Vorhersage der Fahrfaktoren durch die Testfaktoren erstellt, zusätzlich wurde versucht, über eine Regression die Fahrreignung vorherzusagen. Im multivariaten Kontext konnten bei den Regressionsanalysen 67 % der Varianz der gewichteten/ungewichteten Summe der Fahrverhaltensfehler durch die Leistungstestvariablen erklärt werden. Die Pfadanalysen nach Durchführung der Faktorenanalysen lieferten mit normierten Fit-Indices zwischen 0,93 und 0,96 gute Erklärungsmodelle des multivariaten Bedingungsgefüges. Es wurden keine Aussagen über „falsch-positive“ oder „falsch-negative“ Einordnungen gemacht.

TAP-M

Die Untertests der TAP-M wurden vor allem von AKINWUNTAN überprüft (AKINWUNTAN et al., 2006: 68 Patienten nach Schlaganfall; AKINWUNTAN et al., 2002: 104 Patienten nach Schlaganfall), andere Autoren verwendeten nur einzelne Untertests (De RAEDT et al., 2000: Inkompatibilität; HELD et al., 1993: geteilte Aufmerksamkeit). AKINWUNTAN (AKINWUNTAN et al., 2002) konnte für alle Untertests Korrelationen zur Fahrprobe feststellen, in der Folgestudie 2006 wurde jedoch der Untertest Flexibilität nicht mehr überprüft, da er in der Studie 2002 nur eine grenzwertige Vorhersage der Entscheidung fahrg geeignet/nicht fahrg geeignet

erlaubte. Der Autor betont zusätzlich, dass Patienten, die möglicherweise zurzeit noch nicht fahr-tüchtig sind, in neuropsychologischen Testverfahren unauffällig abschneiden können, sich im Straßenverkehr dann aber durch fehlende Kompensationsmöglichkeiten als schlechte Fahrer herausstellen.

Weitere Verfahren

Der UFOV hat bei zwei unterschiedlichen Studien desselben Autors unter gleichen Outcome-Variablen und vergleichbaren Stichproben einmal einen Vorhersagewert zur Fahrprobe gehabt, in der Folgestudie jedoch nicht. In der ersten Untersuchung zeigte sich eine Korrelation von $-0,38$ zur Fahrprobe und von $-0,43$ zur Gruppenentscheidung zur Fahrreignung (dichotomes Urteil, vor allem die Leistung in der Fahrprobe war ausschlaggebend für die Gruppenentscheidung (AKINWUNTAN et al., 2002)). In der Folgestudie konnte sich dieser Zusammenhang nicht bestätigen (weder für die Gruppenentscheidung noch für die Fahrprobe: (AKINWUNTAN et al., 2006)). In einer Studie von De RAEDT hatte der UFOV gute Vorhersageeigenschaften (Korrelation UFOV zu Fahrprobe $r = -0,66$ (De RAEDT et al., 2000)).

Die Rey-Osterrieth-Figur wurde nur von AKINWUNTAN in zwei Studien eingesetzt (AKINWUNTAN et al., 2006; AKINWUNTAN et al., 2002). Während sie in der ersten Studie die höchsten Korrelationen zu Fahrprobe ($r = 0,48$) und Gesamtbewertung ($r = 0,42$) aufwies, zeigte sich in der zweiten Studie nur noch ein Zusammenhang zur Gesamtbewertung ($r = -0,40$), aber nicht mehr zur Fahrprobe ($r = 0,19$).

Der Linienverfolgungstest (LVT) und der Tachistoskopische Verkehrsauffassungstest (TAVT) zeigten sich nicht prädiktiv für die Fahrprobe (HANNEN et al., 1998; HELD et al., 1993). Der TMT wurde nur in einer Studie in Deutschland/Österreich/Belgien eingesetzt (KROLL et al., 2003). Dort korrelierte er mit der Fahrnote, lieferte aber in einer Regression keinen signifikanten Beitrag. Der d2 wurde in zwei deutschen Studien zur Vorhersage der Leistung in einer Fahrprobe eingesetzt, in beiden Studien hatte das Testverfahren keinen Vorhersagewert (HANNEN et al., 1998; HELD et al., 1993).

KARNER et al. (2001) haben in einer Untersuchung mit 127 Teilnehmern die Verfahren Expertensystem Verkehr, ART 90, WDG und WRG miteinander verglichen. Hierbei konnten sie gute Interkorrelationen

der einander entsprechenden Untertests belegen, was für eine weitgehende Äquivalenz der Testverfahren beziehungsweise der damit gemessenen Konstrukte spricht.

In internationalen Studien zeigte sich, dass vor allem der TMT B eine gute Vorhersage der Fahrprobe liefern konnte (GRACE et al., 2005; NOVACK et al., 2006; REGER et al., 2004). Ebenso konnten die Rey-Osterrieth-Figur und der UFOV in Studien überzeugen (GRACE et al., 2005; MARSHALL et al., 2007; NOVACK et al., 2006). Bei Patientengruppen (Demenz, Parkinson) konnte die Minimal State Examination (MMSE) in Metaanalysen und in Originalbeiträgen zum Teil zur Vorhersage der Fahreignung beitragen (ANSTEY, WOOD, LORD & WALKER, 2005; SINGH et al., 2007), zum Teil jedoch nicht (REGER et al., 2004).

Unterschiede zwischen internationalen Studien und den Studien aus Deutschland/Österreich/Belgien können nicht nur davon herrühren, dass die in Deutschland zugelassenen Systeme nicht international eingesetzt werden, sondern können auch von methodischen Unterschieden stammen (siehe oben: Beispiel TMT).

In einer Studie wurden die Testverfahren Expertensystem Verkehr, ART 90, WDG und WRG untereinander verglichen (KARNER et al., 2001). Die Autoren kamen zu dem Schluss, dass die einander entsprechenden (Unter-)Tests unabhängig von der Präsentation dasselbe messen. Ob diese Verfahren einen Vorhersagewert zur Fahreignung liefern, wurde jedoch von den Autoren nicht untersucht.

3.1.4 Fazit – Forschungsergebnisse

Auch bei der spärlichen Datenlage in der Forschung kann davon ausgegangen werden, dass die in Deutschland gängigen Verfahren wie ART 2020, Expertensystem Verkehr (bzw. Wiener Determinationsgerät und Wiener Reaktionsgerät) und TAP-M Aussagen über einen gewissen Zusammenhang mit der Fahreignung erlauben. Eine Standardisierung der eingesetzten Methoden sowie der Fahrprobe (siehe hierzu auch Kapitel 3.4) ist jedoch dringend angezeigt, um die Ergebnisse vergleichbar zu machen und zu wissenschaftlich abgesicherten Ergebnissen zu kommen.

3.2 Angaben der Testhersteller zur Kriteriumsvalidität in den Testmanualen

Neben den in unabhängigen Fachzeitschriften publizierten Forschungsergebnissen geben die Testhersteller in den Testmanualen oftmals weitere Validierungsstudien an. Zum Teil werden jedoch auch Studien angegeben, die ebenfalls in Fachzeitschriften veröffentlicht wurden, es kann also im Folgenden zu Überschneidungen mit vorher dargestellten Ergebnissen kommen.

3.2.1 ART 2020

Zum ART 2020 bestehen umfangreiche Validierungsstudien, die auch separat veröffentlicht (BUKASA et al., 2003a; BUKASA et al., 2001; BUKASA et al., 2003b) und daher bereits unter Kapitel 3.1.3 benannt wurden. In der letzten Validierungsstudie von 2003 wurden die Daten von 123 Probanden ausgewertet. Die Werte der einzelnen Testverfahren wurden faktorenanalytisch untersucht (6 Faktoren), ebenso die Einzelbewertungen des Fahrverhaltens (4 Faktoren).

Die sechs Testverfahrensfaktoren beinhalten:

1. (22 % erklärte Varianz) unterschiedliche Aspekte der Informationsverarbeitungsgeschwindigkeit in ihrer negativen Ausprägung, erfasst bei den wesentlichsten Testverfahren der Fahreignungsdiagnostik (RST3, LL5, DR2, PVT, SET3, Q1, TT15).
2. (11 % erklärte Varianz) sensomotorische Koordination, auch in Verbindung mit der peripheren Wahrnehmung. Der Faktor besteht aus den großen Fehlern im SENSO.
3. (9 % erklärte Varianz) Fehlervariablen unter erhöhter Belastung bzw. Mehrfachanforderungen. Hier gehen Fehler im RST3, PVT und MAT ein.
4. (7 % erklärte Varianz) kleine Fehler im SENSO. Dieser auf die kleinen Fehler eingeschränkte Fehlerfaktor dokumentiert, dass es einen qualitativen Unterschied hinsichtlich Fein- und Grobmotorik gibt.
5. (7 % erklärte Varianz) Überforderung in der Grobmotorik. Anzahl und Dauer „großer“ Fehler sowie die Gegenlenkbewegungen bei erhöhtem festgesetzten Tempo im SENSO bestimmt.

6. (5 % erklärte Varianz) Fehlervariablen aus dem reaktiven, perzeptiven, konzentrativen Intelligenz- und Gedächtnisbereich.

Die vier Fahrverhaltensfaktoren sind:

1. (27 % erklärte Varianz) Gefährdungsaspekte im Fahrverhalten.
2. (15 % erklärte Varianz) Defizite in der Schnelligkeit verkehrsbezogener Informationsaufnahme und in der Genauigkeit der Informationsumsetzung sowie Interaktionsprobleme mit nicht-motorisierten Verkehrsteilnehmern.
3. (13,5 % erklärte Varianz) Schwierigkeiten beim Verlassen der eigenen Fahrbahn und Einordnen in die neue Fahrspur zum passenden Zeitpunkt, Mängel beim rechtzeitigen Signalisieren von Richtungsänderungen.
4. (12,9 % erklärte Varianz) Seitenabstand zu gering.

Mit Hilfe einer Pfadanalyse konnten die Fahrverhaltensfaktoren durch die Testfaktoren vorhergesagt werden. Per Pfadanalyse haben die Autoren ein Modell erstellen können mit einem normierten Fit-Index von 0,93 bzw. einem relativen Fit-Index von 0,80. Der Median der Residuen liegt mit 0,087 unter der obligaten 10 % Schranke, d. h. es bleibt in etwa nur 9 % Residualvarianz unerklärt (BUKASA et al., 2003a).

3.2.2 Expertensystem Verkehr

In den Untertests des Expertensystems Verkehr bestehen leichte Unterschiede in der Art der Validierung. Das Außenkriterium der Fahrprobe wurde bei folgenden Untertests eingesetzt: 2-Hand-Koordination, Adaptiver Matritzen-Test, Determinationstest, Periphere Wahrnehmung, Linienverfolgungstest, Reaktionstest und dem Tachistoskopischen Verkehrsauffassungstest. Bei den genannten Verfahren zeigten sich durchweg signifikante korrelative Zusammenhänge zur Fahrprobe. Da keine Angaben zu Cut-offs gemacht wurden und sich die gemachten Aussagen nur auf korrelative Zusammenhänge stützen, können keine Angaben zur korrekten/inkorrekten Klassifikation von Probanden in fahrgeeignet/nicht fahrgeeignet gemacht werden.

Die Ergebnisse der Vorhersage der Leistung in der Fahrprobe schwanken hierbei. So konnten beispielsweise einige unabhängige Studien keine Zusammenhänge zwischen dem TAVT und einer

Fahrprobe finden (HANNEN et al., 1998.; HELD et al., 1993), die in dem Manual angegebenen Studien konnten jedoch solche Zusammenhänge belegen.

3.2.3 TAP-M

Im Manual des Herstellers werden verschiedene Studien angeführt, die an einer Fahrprobe validiert wurden. Hierbei wurden in der Regel nicht nur die Einzelfaktoren der Tests mit der Fahreignung (ermittelt über verschiedene Testprotokolle, aber immer über eine Fahrprobe) korreliert, sondern auch Faktoren über die Einzelleistungen gebildet, die dann ihrerseits mit der Leistung in der Fahrprobe korreliert wurden. Hier zeigten sich vor allem einheitlich gute Zusammenhangswerte für den Untertest Visuelles Scannen, bei den ermittelten Faktoren waren vor allem Faktoren, die visuell-räumliche Aufmerksamkeit bei Beteiligung von Augenbewegungen zusammenfassten, aussagekräftig. Da auch in diesem Manual nur korrelative Zusammenhänge berichtet werden, kann keine Aussage zur Güte der Klassifikation fahrgeeignet/nicht fahrgeeignet gemacht werden.

3.2.4 Corporal-A

Im Testmanual werden keine Studien angegeben, die an einer Fahrprobe validiert wurden.

3.2.5 d2

Im Testmanual werden keine Studien angegeben, die an einer Fahrprobe validiert wurden.

3.2.6 Konzentrations-Leistungs-Test

Im Testmanual werden keine Studien angegeben, die an einer Fahrprobe validiert wurden.

3.2.7 Zusammenfassung

Gerade bei den Testverfahren, die nicht explizit zur Beurteilung der Fahreignung entwickelt wurden oder deren Einsatz zu diesem Zweck im Manual empfohlen wird, fehlen Angaben zur Validierung am Außenkriterium der Fahrprobe. Bei den anderen Verfahren (ART 2020, Expertensystem Verkehr und TAP-M) werden – mit Ausnahme des Corporal-A – entsprechende Ergebnisse vorgelegt. Diese beziehen sich jedoch im Einzelfall nicht immer auf alle Untertests (Expertensystem Verkehr). Des Weiteren werden keine Angaben zur Klassifikation der Testperson in

fahrg geeignet/nicht fahrg geeignet gemacht, sondern nur korrelative Zusammenhänge angeben.

3.3 Vergleich der Ergebnisse der empirischen Forschung und der Angaben der Testhersteller

Bei den Test-Systemen Expertensystem Verkehr, ART 2020 und TAP-M sind die Ergebnisse der in Zeitschriften veröffentlichten empirischen Forschung weitgehend deckungsgleich mit den Angaben der Manualhersteller. Auch wenn einzelne Untertests uneinheitliche Ergebnisse im Hinblick auf die Vorhersagbarkeit der Fahrprobe liefern, sind sie mit Einschränkungen zur Vorhersage der Fahreignung anwendbar. Zu beachten ist hier allerdings, dass die Angaben in den Manualen zum Teil auf den gleichen Daten beruhen, die in Fachzeitschriften publiziert wurden. Bei dem Testverfahren UFOV fällt zudem auf, dass die umfangreiche Datenlage in Fachzeitschriften (meistens jedoch ohne die Vorhersage einer Fahrprobe und daher hier nicht aufgeführt) vor allem von der Forschungsabteilung des Herstellers mitgetragen wird. Ein Vorteil dieser Veröffentlichungen liegt jedoch darin, dass die Ergebnisse leichter nachvollzogen werden können, als wenn die Daten nicht oder nur unzureichend in den Manualen der Hersteller angegeben werden. Durch die in Fachzeitschriften breitere Darstellung von Versuchsplanung, verwendeten statistischen Methoden und der aus den Ergebnissen gezogenen Schlüsse können leichter die Qualität der Studie und die Aussagekraft der Ergebnisse beurteilt werden.

3.4 Exkurs: das Kriterium der Fahrprobe

Bei Durchsicht der Literatur fiel auf, dass das Außenkriterium der Fahrprobe sehr unterschiedlich realisiert wird. Dies wirft zum einen das Problem auf, dass unterschiedliche Studien mit unterschiedlich durchgeführten Fahrproben schwer vergleichbar sind. Zum anderen ist fraglich, ob jede Art der Durchführung einer Fahrprobe als Kriterium ausreichend ist. Sinnvoll ist es, sich zumindest an den Anforderungen der 3. EU-Führerschein-Richtlinie (Richtlinie 2006/126/EG des Europäischen Parlaments und des Rates vom 20. Dezember 2006 bzw. deren Vorläufern) zu orientieren. Hier heißt es: „Prüfungsdauer und Prüfungstrecke müssen so bemessen sein, dass die Fähigkeiten und Verhaltensweise gemäß dem Kapitel B dieses Anhangs beur-

teilt werden können. Die Mindestfahrzeit [...] darf in keinem Falle weniger als 25 Minuten für die Klassen A, A1, A2, B, B1 und BE [...] betragen.“ Weiter wird auf den Ort der Prüfungsleistung verwiesen, die Prüfung findet nach Möglichkeit „außerhalb geschlossener Ortschaften, auf Schnellstraßen und auf Autobahnen (oder ähnlichen Straßen) sowie auf allen Arten von Straßen in bebautem Gebiet statt (Wohngebiete, Gebiete mit Beschränkung auf 30 km/h und 50 km/h, städtische Schnellstraßen) mit den verschiedenartigen Schwierigkeiten, auf die ein Fahrer stoßen kann“. Verschiedene Verkehrs-dichten sollten ebenfalls realisiert werden. Auch an den Fahrprüfer werden umfangreiche Anforderungen gestellt, die allerdings im Falle von Fahrlehrern als Prüfer (BUKASA et al., 2003a; KROLL et al., 2003; SOMMER et al., 2006) als gegeben angesehen werden können.

Es werden also Festlegungen gemacht in Bezug auf Dauer und Strecke der Fahrprobe, Inhalte der Bewertungskategorien sowie bezüglich der Qualifikation des Prüfers.

Während die in den Studien dokumentierte zeitliche Dauer der Fahrprobe (gemessen an der EU-Richtlinie 2006/126/EG und 2000/56/EC mit einem angegebenen Minimum von 25 Minuten) als ausreichend zu beurteilen ist (mit Ausnahme von SINGH et al., 2007 mit 15 Minuten) und auch die Orte der Fahrprobe vergleichbar sein dürften (die gewählten Strecken orientieren sich an dem Ausführungsort der Studie, es handelt sich jedoch immer um Fahrten im realen Straßenverkehr mit einem vergleichbaren Anteil an beruhigten Zonen, normalem Stadtverkehr und Autobahnfahrten, die, sofern angegeben, auch in etwa gleichen Verhältnissen gewählt wurden), bestehen starke Differenzen in der Ausbildung bzw. dem fachlichen Hintergrund der Bewerter (Fahrlehrer vs. unabhängiger Rater oder beide getrennt voneinander) sowie in der Bewertung der Fahrprobe. Die Bewertung der Fahrprobe erfolgt in den Studien in der Regel über Skalen, welche von dichotomen Aussagen (bestanden, nicht bestanden) zu mehrstufigen Skalen reichen. Bei den mehrstufigen Skalen können 3 Stufen beurteilt werden (z. B. „sicher“, „mit Einschränkungen“, „unsicher“ (BROWN et al., 2005)), es werden Schulnoten vergeben (z. B. HELD et al., 1993), zum Teil mit späterer Dichotomisierung (1-4 bestanden, 5-6 durchgefallen, z. B. HANNEN et al., 1998), es werden Scores von 0 - 108 vergeben (Washington University Road Test, z. B. GRACE et al., 2005) oder der Prozentrang richtiger Verhaltensweisen beur-

teilt (KROLL et al., 2003). Hier muss erörtert werden, ob mehrstufige Skalen sinnvoll sind oder ob es reicht, mittels psychometrischer Verfahren die Fahreignung im Sinne eines dichotomen Urteils 'fahrgeeignet/nicht fahrgeeignet' vorherzusagen. Grenzbereiche zwischen den beiden Einschätzungen müssten dann durch eine Fahrprobe abgeklärt werden, bei eindeutigen Testergebnissen (und einer entsprechenden Vorhersagegenauigkeit der Testverfahren!) könnte auf die Fahrprobe verzichtet werden. Die Angabe von Konfidenzintervallen oder wahlweise eine Abstufung von Cut-Offs im Sinne von „fahrgeeignet“/„möglicherweise fahrgeeignet: Fahrprobe erforderlich“/„nicht fahrgeeignet“ wäre in der Praxis die hilfreichste Methode. So könnte dem Testanwender eine Hilfestellung gegeben werden, bei welchen Probanden eine nachgeschaltete Fahrprobe nötig ist und welche Probanden sicher klassifiziert werden können.

Neben der inhaltlichen Fragestellung um die Zielgröße der Fahrprobe besteht noch eine methodische Problematik: Mit der Skalenqualität der Fahrerurteilung ändern sich auch die einsetzbaren statistischen Verfahren, was wiederum die Vergleichbarkeit verschiedener Studien herabsetzt.

3.5 Zusammenfassung

Aufgrund der stark heterogenen Verwendung von miteinander kombinierten Testverfahren, Studienpopulationen und Stichprobenumfängen, statistischen Methoden und schlussendlich stark verschiedenen Durchführungsweisen der Fahrprobe ist eine exakte wissenschaftliche Beurteilung der Testverfahren im Hinblick auf die Kriteriumsvalidität „Fahreignung“ derzeit nicht möglich. Die in Fachzeitschriften veröffentlichten Forschungsergebnisse und die Herstellerangaben lassen jedoch den Schluss zu, dass vor allem die Testbatterien Expertensystem Verkehr, ART 2020, TAP-M, die in Deutschland zur Vorhersage der Fahreignung in Gebrauch sind bzw. zu diesem Zweck entwickelt wurden, einen gewissen Vorhersagewert zur eigentlichen Fahreignung erbringen können. In internationalen Studien haben sich auch TMT und UFOV bewähren können, hier stehen Vergleiche zum Zusatznutzen bzw. zur Überlegenheit gegenüber den oben genannten Testverfahren noch aus.

Zu dem mittlerweile in allen Bundesländern Deutschlands zugelassenen Testverfahren Corporal-A konnten bis Mitte Juni 2007 keine Studien ge-

funden werden, die den Vorhersagewert zur Fahrprobe behandelt hätten. Auch im Manual sind keine Angaben zu einer Fahrverhaltensvalidierung gemacht. Somit ist der Corporal-A der einzige zugelassene Test, zu dem derzeit offensichtlich weder in der Literatur noch im Manual Angaben zur Kriteriumsvalidität vorliegen.

4 Methodisches Vorgehen

Zur Erreichung des Forschungszieles wurden folgende, aufeinander abgestimmte methodische Vorgehensweisen gewählt (die Einzelergebnisse werden in den nächsten Kapiteln dargestellt):

4.1 Schriftliche Befragung von Gutachtern über die in der Praxis eingesetzten Tests

Durch eine repräsentative schriftliche Befragung von Gutachtern sollten umfassende Informationen zur Anwendung der in der Praxis eingesetzten Tests erfasst werden (Häufigkeiten und bestehende Problembereiche, Kapitel 5).

4.2 Auswertung der Testmanuale der erfragten und recherchierten Tests hinsichtlich Standardisierung

Die Testmanuale der erfragten Tests wurden dahingehend überprüft, ob sie die in der FeV geforderten Kriterien erfüllen (eindeutige Zuordnung zu den in der FeV genannten kognitiven Bereichen, wissenschaftliche Standardisierung, Validierung unter Aspekten der Verkehrssicherheit).

Hierzu wurde in enger Abstimmung mit der BAST eine wissenschaftliche Kriterienliste erarbeitet, mit der die wesentlichen Bestandteile des jeweiligen Tests (postulierte psychische Dimension, Eichstichprobe, Objektivität der Auswertung, Validierung etc.) erfasst und unter Anwendung der geforderten Kriterien beurteilt wurden (Kapitel 6.1). Parallel wurden Validierungsstudien der einzelnen Tests recherchiert, sofern sie allgemein zugänglich veröffentlicht sind (siehe Kapitel 3). Auf Basis dieser Informationen wurde die von der FeV geforderte wissenschaftliche Güte des jeweiligen Tests insgesamt überprüft (Kapitel 7).

4.3 Einteilung der Tests in drei Kategorien

Die recherchierten Tests wurden aus theoretischer und praktischer Sicht anschließend in drei Kategorien eingeteilt (Kapitel 7.9 und 7.10):

Gruppe 1: Tests, die thematisch den in der FeV genannten Bereichen entsprechen, die jedoch nicht entsprechend normiert wurden und für die keine Validierungsstudien vorliegen.

Gruppe 2: Tests, die thematisch den in der FeV genannten Bereichen entsprechen und darüber hinaus auch entsprechend normiert sind, für die jedoch keine bzw. unzureichende Validierungsstudien vorliegen.

Gruppe 3: Tests, die den in der FeV genannten Kriterien entsprechen.

4.4 Diskussion der Ergebnisse in einer moderierten Expertengruppe

Bei einem Expertentreffen wurden die zentralen methodischen Fragen zur Fahreignungsdiagnostik diskutiert. Ebenfalls thematisiert wurde, ob die in der FeV genannten Bereiche noch derzeitigen wissenschaftlichen Modellen der Aufmerksamkeit entsprechen sowie die zentralen Ergebnisse der schriftlichen Befragung und der Analyse der Testsysteme. Die Auswahl der externen Experten erfolgte durch die BAST.

4.5 Empfehlungen zur Anwendung der Begutachtungs-Leitlinien unter methodischen und neuropsychologischen Aspekten/ Zusammenstellung angemessener Testverfahren

Aufbauend auf den Resultaten der wissenschaftlichen Beurteilung des bestehenden Testkanons auf Basis der von der FeV geforderten Kriterien und der Ergebnisse der Expertenbefragung werden Empfehlungen zur Anwendung der „Begutachtungs-Leitlinien zur Krafftfahreignung“ entwickelt (Kapitel 9.3 bis 11). Dabei werden Fragen der Standardisierung, der Prozenträge und einer möglichen Kompensation erörtert. Abschließend werden dieje-

nigen Messverfahren zusammengestellt, die den Kriterien der FeV und den Begutachtungsleitlinien am besten entsprechen und auch unter ökonomischen Aspekten sinnvoll eingesetzt werden können. Sollte sich als ein Ergebnis zeigen, dass derzeit aus methodischer und/oder neuropsychologischer Sicht der ermittelte Testkanon nicht den von der FeV geforderten Kriterien entspricht, werden Empfehlungen zur Konstruktion neuer Tests bzw. verbesserter Normierungen bereits bestehender Tests gegeben.

5 Befragung von akkreditierten Begutachtungsstellen und ärztlichen Gutachtern

5.1 Fragebogenentwicklung

Um Aussagen über die in der Praxis eingesetzten psychometrischen Testverfahren zur Fahreignung (Häufigkeit des Einsatzes verschiedener Testsysteme, auftretende Probleme, inhaltliche Lücken aus Sicht der Gutachter usw.) machen zu können, wurden zunächst zwei Fragebögen entwickelt: ein Fragebogen für ärztliche Gutachter der Fahreignung sowie ein Fragebogen für die in akkreditierten Begutachtungsstellen tätigen Gutachter der Fahreignung. Die Fragebögen wurden im Vorfeld mit der BAST abgestimmt und können dort eingesehen werden. Inhaltlich unterscheiden sich die Fragebögen kaum, lediglich bei den allgemeinen Angaben gibt es geringfügige Unterschiede. Bei den Ärzten wurde im allgemeinen Teil erfragt, welche Zusatzqualifikation sie besitzen, ob sie selbstständig oder in einer Organisation arbeiten und wie viele Gutachten (mit psychometrisch ausgerichteter Fragestellung) sie im letzten Jahr erstellt haben und wer hierfür meist Auftraggeber war.

Von den in Begutachtungsstellen tätigen Gutachtern sollte im allgemeinen Teil des Fragebogens angegeben werden, wie viele Gutachter in der jeweiligen Begutachtungsstelle arbeiten und wie viele Gutachten im letzten Jahr etwa erstellt wurden.

Anhand der genannten Zahlen sollte abgeschätzt werden, welches Gewicht einzelne Testverfahren in der Praxis haben. Wird ein bestimmter Test beispielsweise von einer größeren Organisation als Standardtest eingesetzt, mit dem jährlich mehrere tausend Gutachten erstellt werden, kann von einem großen Gewicht dieses Testverfahrens in der Praxis ausgegangen werden.

Im zweiten Teil des Fragebogens wird erfragt, welche Tests (Testname und Version) bei einer Überprüfung der Bereiche „Belastbarkeit“, „Orientierungsleistung“, „Konzentrationsleistung“, „Aufmerksamkeitsleistung“ und „Reaktionsfähigkeit“ jeweils eingesetzt werden und wie oft die Instrumente zum Einsatz kommen (siebenstufige Skala von „nie“ bis „immer“).

Im dritten Teil wird erfragt, wie Kraftfahrer einzelne Bereiche kompensieren können, wenn sie den geforderten Prozentrang nicht erreichen konnten, und welche Prinzipien/Kriterien in der jeweiligen Begutachtungsstelle einer Kompensationsprüfung zugrunde liegen. Abschließend wird in zwei offenen Fragen erfragt, ob aus Sicht des Gutachters das bestehende Testinstrumentarium ausreichend ist, um die in der FeV genannten Bereiche zu überprüfen, und ob aus Sicht des Gutachters weitere Problemschwerpunkte bestehen. Besonders durch die Antworten auf die letzten beiden Fragen sollte die spezielle Perspektive der in der Praxis tätigen Einzelgutachter berücksichtigt werden.

5.2 Durchführung der schriftlichen Befragung

Die Adressen für die Befragung der ärztlichen Gutachter wurden im Internet bei den Ärztekammern recherchiert. Angeschrieben wurden solche Ärzte, die auf den Internetseiten der Ärztekammern mit einer entsprechenden Zusatzqualifikation ausgewiesen waren. Dabei handelt es sich weder um eine Zufallsstichprobe noch um eine repräsentative Stichprobe (nicht alle Ärztekammern stellen entsprechende Informationen auf den Internetseiten bereit). Trotzdem darf davon ausgegangen werden, dass hierdurch ein hinreichender Überblick über die in der ärztlichen Begutachtungspraxis eingesetzten psychometrischen Verfahren gewonnen werden konnte. Insgesamt wurden 254 Mediziner angeschrieben: Ärztekammer Nordrhein (166 Adressen), Ärztekammer Holstein (32 Adressen), Ärztekammer Berlin (11 Adressen), Ärztekammer Niedersachsen (11 Adressen), Ärztekammer Westfalen-Lippe (4 Adressen), Ärztekammer Saarland (1 Adresse), Ärztekammer Bayern (13 Adressen), Ärztekammer Sachsen (11 Adressen) und Ärztekammer Hamburg (5 Adressen).

Bei den akkreditierten Begutachtungsstellen sollten möglichst viele Begutachtungsstellen erreicht und möglichst eine repräsentative Stichprobe be-

fragt werden, da angenommen wurde, dass von diesen Begutachtungsstellen der Hauptteil der bundesdeutschen Begutachtungen im Rahmen der psychometrischen Leistungsmessung bei der Fahreignung durchgeführt wird. Grundlage war die im März 2006 aktuelle Liste der BAST von den in Deutschland akkreditierten Trägern der Begutachtungsstellen. Die Adressen der Filialstellen der akkreditierten Träger wurden ebenfalls im Internet auf den Homepages der jeweiligen Organisationen recherchiert. Von den derzeit existierenden etwa 200 Begutachtungsstellen (Stand Mai 2006) wurden 142 Niederlassungen/Standorte direkt angeschrieben, die DEKRA und der TÜV Hessen wurden zentral angeschrieben.

Dem Anschreiben des Forschungsinstituts IfADO wurde ein Begleitschreiben der BAST beigelegt, ebenso ein Fragebogen und ein Rückumschlag, der von den Rücksendern nicht frankiert werden musste. Die Briefe mit den Fragebögen wurden am 16.03.06 versandt, die Rücksendung sollte bis zum 15.04.06 erfolgen.

5.3 Rücklaufquote und Stichprobenbeschreibung

Der Rücklauf bei den angeschriebenen Ärzten ist sehr gering geblieben. Von 225 Ärzten gab es lediglich 27 Antworten (12,0 %). 18 Antworten (8 %) waren nicht verwertbar, weil die angeschriebenen Ärzte keine psychometrischen Leistungsprüfungen durchführen; insgesamt 9 (4,0 %) der Antworten von Ärzten sind verwertbar.

Auch bei den Begutachtungsstellen ist die Rücklaufquote nur unwesentlich höher: Von 143 angeschriebenen Begutachtungsstellen liegen Informationen zu 28 Standorten (20,3 %) vor, von denen insgesamt 25 (18,2 %) verwertbar sind. Trotz Zusage gingen von einigen Organisationen keine Fragebögen ein.

Aufgrund der Umstände entstand der Eindruck, dass in vielen Fällen Begutachtungsstellen zentral angewiesen wurden, keine Rückmeldungen für das Forschungsprojekt zu senden, sodass trotz der avisierten repräsentativen Stichprobe insgesamt nur wenige Rückmeldungen zum Einsatz psychometrischer Verfahren und deren Problemen bei der Anwendung zur Prüfung der Fahreignung vorliegen.

5.4 Auswertung

Aufgrund der schlechten Rücklaufquoten von beiden angeschriebenen Gruppen (Begutachtungsstellen/Ärzte) ist bei der Auswertung zu beachten, dass die Ergebnisse bestenfalls als Tendenz betrachtet werden können. Verlässliche Aussagen, die für beide Gruppen gelten, sind auf Basis der Befragungsergebnisse nicht möglich.

Von den akkreditierten Begutachtungsstellen wurden teilweise Sammelantworten zu mehreren Standorten zurückgesandt, sodass eine Aufschlüsselung der Daten nach Einzelstandorten nicht möglich ist.

Wie erwartet, werden von den befragten akkreditierten Begutachtungsstellen wesentlich mehr Gutachten im Jahr durchgeführt als von einzelnen Ärzten. Das gilt für die absoluten wie auch relativen Häufigkeiten. Während von den einzelnen akkreditierten Begutachtungsstellen im letzten Jahr 580 Gutachten im Durchschnitt durchgeführt wurden (15.065 insgesamt¹), waren es bei den antwortenden Ärzten durchschnittlich 28 Gutachten (254 insgesamt, max. 95 Gutachten, min. 2 Gutachten je Arzt).

5.4.1 Von den akkreditierten Begutachtungsstellen eingesetzte Verfahren

Von den akkreditierten Begutachtungsstellen werden ausschließlich die „großen“ Testsysteme eingesetzt, die von den einzelnen Bundesländern zugelassen und bei der BAST akkreditiert sein müssen. Die Testsysteme wurden in folgender Reihenfolge genannt:

„Wiener Testsystem“ in etwa 12.000 der erhobenen Fälle, der „Art 2020“ in etwa 2.000 Fällen und der „Corporal-A“ in knapp 1.000 Fällen. Ob dieses Bild repräsentativ ist, kann wegen der insgesamt geringen Rückmeldung nicht entschieden werden.

5.4.2 Von Ärzten eingesetzte psychometrische Begutachtungsverfahren

Bei den von Ärzten angewandten Begutachtungsverfahren zeigt sich trotz der kleinen Stichprobengröße ein ähnliches Bild: Bei den 254 rückgemel-

deten Gutachtenfällen wurde in über 220 Fällen das „Wiener Testsystem“ eingesetzt. Übrige Testverfahren kommen bei den antwortenden Ärzten nur selten zum Einsatz.

Selten wird der d2 (vor allem für Konzentrationsfähigkeit/Aufmerksamkeit) eingesetzt. Alle anderen genannten Verfahren wurden bei weniger als 10 von den über 250 Begutachtungen eingesetzt.

5.4.3 Offene Fragen

Bei den offenen Fragen zu Problemen bei der Testdurchführung wurden nur wenige Problemschwerpunkte genannt. Grundsätzlich wird das bestehende Testinstrumentarium von den Gutachtern als ausreichend betrachtet. Bemängelt werden jedoch durch Anfangsnervosität verfälschte Testergebnisse (vor allem bei zunehmendem Alter der Probanden). Auch wird bemängelt, dass Tests geübt werden können und hierdurch das Ergebnis verbessert werden kann.

Es wird ebenfalls bemängelt, dass es keine genauen Vorgaben darüber gibt, wie Kompensationsleistungen zu bewerten sind, wenn ein Proband in einem Test besonders gut abscheidet, in einem anderen jedoch nicht den geforderten Prozentrang erreichen kann. In der Praxis kann durch einfache Messwiederholung ein schlechter Prozentrang verbessert werden (unter Umständen so lange, bis ein „guter“ Prozentrang erreicht wird), vor allem dann, wenn der Eindruck der Nervosität entstanden ist.

Sprachprobleme können die Durchführung von Tests ebenfalls erschweren.

5.4.4 Zusammenfassung der Auswertung

Die Befragung der akkreditierten Begutachtungsstellen und der in diesem Bereich tätigen Mediziner hat nicht den erhofften Überblick verschafft. Aufgrund der geringen Anzahl der Rückmeldungen sind die Ergebnisse der Auswertung nur unter großen Einschränkungen zu interpretieren.

Zur Begutachtung wird demnach vor allem das Wiener Testsystem eingesetzt. Änderungswünsche zum bestehenden Testinstrumentarium wurden praktisch nicht geäußert. Es wurden jedoch Hinweise gegeben, dass die Vorschriften zur Interpretation der Messergebnisse (Kompensationsmöglichkeiten) derzeit nicht ausreichend sind, da klare Kriterien hierzu fehlen.

¹ Min.-Max.-Angabe nicht möglich (Sammelantwort)

Bei der Durchführung der Testverfahren scheint die Anfangsnervosität besonders bei älteren Probanden zur Verfälschung von Messergebnissen führen zu können.

6 Beurteilung psychometrischer Tests zur Prüfung der Fahr-eignung

6.1 Entwicklung einer Kriterienliste zur Beurteilung einzelner Tests

Parallel zur Befragung der Gutachter wurde eine Kriterienliste zur objektiven Beurteilung psychometrischer Testverfahren, die in der Eignungsdiagnostik eingesetzt werden, in enger Absprache mit der BAST entwickelt. Grundsätzlich galt die Vorgabe, an die zu untersuchenden Tests die gleichen Kriterien anzulegen, wie sie dem guten wissenschaftlichen Standard der „scientific community“ entsprechen. Hier wurde bei der Entwicklung der abzufragenden Kriterien auf zwei methodische Standardwerke der Psychologie zurückgegriffen (BORTZ & DÖRING, 2003; LIENERT & RAATZ, 1998), außerdem auf die DIN 33430, in der die „Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen“ beschrieben werden (DIN, 2002). Die Kriterienliste wurde auf Basis dieser drei Standardwerke erstellt.

Basis für dieses Vorgehen sind die in der FeV geforderten Kriterien „wissenschaftliche Standardisierung“ und „Validierung unter Aspekten der Verkehrssicherheit“ für die in der Begutachtung eingesetzten psychometrischen Testverfahren. Ebenso wird das in den Begutachtungs-Leitlinien zur Kraftfahrereignung zugrunde gelegte Konzept der Prozenträge überprüft.

Die entwickelte Kriterienliste (die vollständige Liste kann bei der BAST eingesehen werden) zur Beurteilung der Testverfahren gliedert sich in sechs Bereiche:

Allgemeine Angaben

Hier werden Testname, Kürzel des Namens und die Versionsnummer erfasst.

Anforderungsbereiche nach FeV, Anlage 5

Im nächsten Kapitel werden die Anforderungsbereiche nach FeV, Anlage 5, dargestellt. Hier wer-

den auf Basis der Angaben im Testmanual die Zielsetzung und der Anwendungsbereich sowie die psychologischen Dimensionen dokumentiert, die der Test nach Ansicht der Autoren messen soll („Belastbarkeit“, „Orientierungsleistung“, „Konzentrationsleistung“, „Aufmerksamkeitsleistung“ und „Reaktionsfähigkeit“). Danach werden die von den Testautoren genannten Angaben zur Definition der Verhaltensbereiche, den postulierten psychologischen Modellen und den theoretischen Grundlagen des zu erfassenden Verhaltensbereiches und deren Operationalisierung dargestellt (z. B. Wie wurden die genannten psychischen Funktionen operationalisiert? Auf welche Art und Weise soll das theoretische Konstrukt gemessen werden?). Auch wird der Frage nachgegangen, ob es sich um eine eindeutige oder um eine mehrdeutige Messung handelt (Werden die genannten psychischen Funktionen mit bestimmten Subtests oder mit dem gesamten Testverfahren erfasst? Wird eine psychische Funktion mit mehr als einem Subtest gemessen? Werden mit einem Subtest verschiedene Verhaltensbereiche erfasst?). In den nächsten Kapiteln werden bei der Kriterienliste die grundsätzlichen, testtheoretisch bedeutsamen Kriterien erfasst und kritisch dargestellt:

Objektivität

Hier wird dokumentiert, wie gut die Durchführungs-, Auswertungs- und Interpretationsobjektivität des Tests einzuschätzen ist und wie eindeutig sie ist (z. B. Vorhandensein einer Auswertungsschablone).

Zuverlässigkeit (Reliabilität)

Es wird beurteilt, wie hoch die Reliabilität insgesamt einzuschätzen ist und ob die von den Testautoren gemachten Angaben hierzu ausreichend und aktuell genug sind (z. B. Angaben zur internen Konsistenz, Nachvollziehbarkeit der Reliabilitätsberechnungen usw.).

Für die Berechnung der Reliabilitätsindici ist es ebenfalls von Bedeutung, ob die Reliabilitätsberechnungen mit den Daten der Eichstichprobe durchgeführt wurden oder andere Stichproben bei den Reliabilitätsberechnungen zugrunde lagen.

Aspekte, die es bei der Reliabilitätsschätzung bei computergestützten Reaktionstests im Besonderen zu beurteilen gilt, sind Kapitel 6.2 zu entnehmen.

Gültigkeit (Validierung unter Aspekten der Verkehrssicherheit)

Hier werden die Angaben zu den verschiedenen Gültigkeitskriterien des Tests analysiert (z. B. Korrelationen mit anderen Testverfahren). Das sind vor allem Angaben über verkehrsrelevante Gültigkeitskriterien und die Art dieser Außenkriterien sowie Angaben zur Konstruktvalidität, also zur Frage, wie gut die zu messenden postulierten Konstrukte mit psychologischen Modellen übereinstimmen. Ebenso gibt es bei diesem Punkt die Möglichkeit, wissenschaftliche Studien zum Test zu dokumentieren.

Normwerte nach FeV

Hier wird überprüft, ob die von den Begutachtungs-Leitlinien zur Krafftahreignung geforderten Einteilungen in Prozenträngen vorliegen, um auf Basis des Tests die entsprechenden Normwerte einhalten zu können.

In diesem Kapitel wird auch der grundsätzlichen Frage nachgegangen, auf welcher Population die Eichstichprobe des Tests beruht. Die ermittelten Normen können für bundesdeutsche Autofahrer und Autofahrerinnen nur dann gelten, wenn die Eichstichprobe eine Zufallsauswahl der Population darstellt, für die sie später gelten soll.

6.2 Anmerkungen zur Reliabilitäts-schätzung bei computergestützten Reaktionstests

Eine Problematik bei der Bestimmung der Reliabilität von Reaktionstests mit kritischen und nicht-kritischen Reizen ergibt sich aus dem Auftreten von Fehlreaktionen und Auslassungen bei einem Teil der Probandenstichprobe, anhand derer die Messgenauigkeit berechnet werden soll. Dies führt dazu, dass beispielsweise Reaktionszeitreihen länger werden, wenn zusätzliche falsche Alarmer (Fehlreaktionen) seitens des Probanden erfolgen. Reaktionszeit X des Probanden A entspräche in einem solchen Fall nicht Reaktionszeit X bei Proband B, wenn dieser bei dem Durchgang X-5 (z. B. einem nicht-kritischen Reiz) fälschlicherweise reagiert hätte. Dies schließt jedoch eine Berechnung der internen Konsistenz, wenn z. B. der Test in so viele Teile aufgesplittet werden soll, wie er Items/kritische Reize aufweist, noch nicht prinzipiell aus, sofern man sich lediglich auf die validen Reaktionen beschränkt.

Schwerwiegender ist das (je nach Schwierigkeitsgrad des Testverfahrens durchaus häufigere) Auftreten von Auslassungen kritischer Reize. Dies führt dazu, dass der Datensatz einer entsprechenden Person mit zumindest einer Auslassung nicht mit zur Berechnung der internen Konsistenz herangezogen werden kann, da für die Berechnung sämtlicher Item-Interkorrelationen vollständige Datensätze vorausgesetzt werden. Auf diese Weise kann sich die effektive Gesamtgröße einer Probandenstichprobe erheblich reduzieren. Dies kann nur durch die Verwendung von entsprechenden Korrekturverfahren, bei denen fehlende Werte durch andere Werte (z. B. Mittelwert der Reaktionszeiten der Probanden über die restlichen Items) ersetzt werden (Missing-Value-Korrektur), vermieden werden. Je häufiger solche Ersetzungen in einer Stichprobe jedoch gemacht werden müssen, desto ungenauer ist der resultierende Reliabilitätskoeffizient.

Die Bestimmung der Reliabilität nach der Testhalbierungsmethode ist von diesen Problemen nicht betroffen. Hierbei werden pro Testhälfte Kennwerte (z. B. Mittelwert, Median, Anzahl der Auslassungen etc.) berechnet, woraufhin beide Testhälften auf der Basis dieser Kennwerte korreliert werden und mittels der Spearman-Brown-Formel zur Testverlängerung die Reliabilität des Gesamttests geschätzt wird. Einzelne Auslassungen wirken sich dann nicht oder nur wenig aus, sofern noch genügend valide Reaktionen zur Berechnung der Kennwerte vorliegen. Auch die Bestimmung der Retestreliabilität leidet nicht an den o. g. Einschränkungen.

Wenn also computergestützte Testverfahren, bei denen sowohl kritische als auch nicht-kritische Reize auftreten, Reliabilitätsschätzungen zur internen Konsistenz (z. B. Cronbach Alpha, Kuder-Richardson) angeben, ist zusätzlich zu fordern, dass die Anzahl der Probanden, deren Daten effektiv verwendet wurden, angegeben wird. Falls die Daten aller Probanden Eingang fanden, sollte das Verfahren der Missing-Value-Korrektur angegeben werden.

6.3 Vorgehensweise bei der Testanalyse

Der Kriterienkatalog zur Testanalyse soll ein objektives Beurteilungsschema bieten und enthält deshalb weitgehend standardisierte Antwortformate für die Beurteilung der einzelnen Testgütekriterien. Freie Antwortfelder für Anmerkungen am Ende jedes Gütekriteriums boten die Möglichkeit eines abschließenden Urteils zu jedem Kriterium auf der

Basis der vorangegangenen standardisierten Beurteilungen.

Jedes der interessierenden Verfahren wurde einzeln anhand des Kriterienkataloges beurteilt. Bei den untersuchten Testsystemen, die aus mehreren Untertests bestehen, wurde jeder Untertest einzeln der Testanalyse auf Basis der entwickelten Kriterienliste unterzogen. Grundlegend für die Testbewertung ist dabei das Testmanual; für das Vorgehen bei der Analyse die Verfahrenshinweise, die das Testkuratorium der Förderung deutscher Psychologenverbände (Testkuratorium der Förderung deutscher Psychologenverbände, 1986) und die DIN 33430 vorsehen (DIN, 2002; Testkuratorium der Förderung deutscher Psychologenverbände, 1986).

7 Ergebnis der Testanalysen

Dieses Kapitel beschreibt Arbeiten zur Analyse der in der Vorbefragung ermittelten, von akkreditierten Begutachtungsstellen und ärztlichen Gutachtern eingesetzten Tests und Verfahren zur Überprüfung der Fahreignung. Da es im Forschungsprojekt vor allem darum gehen soll, bestehende Probleme aufzuzeigen und Optimierungsbedarf zu formulieren, werden die Ergebnisse der Testanalysen in anonymisierter Form dargestellt. An den inhaltlichen Aussagen ändert sich hierdurch selbstverständlich nichts. Zunächst werden die Vorgehensweise bei der Beschaffung der Testmanuale und die Resonanz der Verlage beschrieben. Anschließend wer-

den die Ergebnisse der einzelnen Testanalysen zusammenfassend dargestellt. Dabei gliedert sich das Kapitel nach den Testbatterien A, B, C, D und anschließend nach den Tests E und F. Die zusammenfassenden Beurteilungen der Testverfahren basieren auf den Testanalysen der einzelnen Tests und liegen der BAST vor. Der Fokus der zusammenfassenden Beurteilung liegt demzufolge, wie bei den einzelnen Testanalysen selbst, auf den zuvor festgelegten Testgütekriterien, die unter Kapitel 6.1 dargestellt sind (zum Aufbau der Kriterienliste siehe Kapitel 6.1).

7.1 Beschaffung der Testmanuale

Zunächst erfolgt die Beschreibung der Beschaffung der Testverfahren. Aufbauend auf den Ergebnissen der Befragung von akkreditierten Begutachtungsstellen und ärztlichen Gutachtern wurden die Tests für den nächsten Schritt der Testanalyse festgelegt. Dabei wurden vor allem Tests, die von der BAST (und den Bundesländern) zugelassen sind und von den akkreditierten Begutachtungsstellen verwendet werden, berücksichtigt.

Anschließend wurden die entsprechenden Verlage bzw. Bezugsquellen recherchiert und über die BAST angeschrieben. Unter Angabe des Projekts, des damit betrauten Projektträgers und einer ausführlichen Beschreibung des Projektzieles wurden die Verlage gebeten, die interessierenden Manuale und Testinformationen bereitzustellen. Tabelle 3 enthält eine Übersicht der Verlage bzw. Testent-

Verlag und Bezugsort	Test	
Dr. Schuhfried GmbH	Wiener Testsystem	Enthält u. a. LVT, TAVT, Determinationsgerät, Reaktionsgerät, Cognitronem (S1) „COG“
	Verkehrspsychologische Testbatterie	Enthält u. a. LVT, TAVT, Determinationsgerät, Reaktionsgerät, Cognitronem (S1) „COG“
	PASAT	Progressiver Auditiver Serieller Additions-Test
Harcourt Test Services	INKA	Inventar Komplexer Aufmerksamkeit
	PASAT	Progressiver Auditiver Serieller Additions-Test
HASOMED GmbH	CombiTest	
Institut für Testentwicklung und -anwendung – itea	Corporal-A	Enthält u. a. DgLok, DgOr, InterLok
Kuratorium für Verkehrssicherheit – Wien	ART 2020 ART 90	Enthält u. a. TT 15, Q1, RST 3
Psytest Psychologische Testsysteme	TAP-M	Testbatterie zur Aufmerksamkeitsprüfung – Version Mobilität
Testzentrale Göttingen	d2	Aufmerksamkeits-Belastungs-Test, 9. überarbeitete und neu normierte Auflage
	KLT	Konzentrations-Leistungs-Test
	INKA	Inventar Komplexer Aufmerksamkeit

Tab. 3: Übersicht der Verlage bzw. Testentwickler und der angeforderten Testmanuale

wickler und angeforderten Testmanuale. Mit einer Ausnahme kamen alle Verlage der Manualanfrage durch die BASt nach.

In einem nächsten Schritt wurde nach Sichtung der zugesandten Unterlagen der Umfang der analysierten Tests eingeschränkt. Dies geschah zum einen aufgrund der zu vernachlässigenden Relevanz einiger Tests für die Überprüfung der Fahreignung. Hiervon betroffen waren die Tests INKA und PASAT. Sie wurden zwar vereinzelt von Medizinern empfohlen, doch wie die Auswertung ergab (s. Kapitel 5.4), kamen sie bei den meisten Gutachten nicht zum Einsatz. Zum anderen wurden fortan die Testsysteme ART 90 und CombiTest außer Acht gelassen, da die Verlage bzw. Testhersteller mitteilten, dass diese Testsysteme nicht mehr vertrieben würden und auch kein Support mehr bestünde.

7.2 Testsystem A

Das Testsystem umfasst mehrere Untertests.

Die zusammenfassende Beurteilung der Testbatterie A basiert auf den Testanalysen zu den darin enthaltenen Einzeltests und der Handlungsanweisung des Testsystems.

Es gilt festzustellen, dass keine einheitliche theoretische Grundlage, die der gesamten Testbatterie zugrunde liegt, zu finden ist. Die übergreifende Handanweisung zum Testsystem enthält vor allem Benutzungshinweise für Anwender und technische Informationen.

Nur bei vereinzelt Tests werden theoretische Grundlagen ausgeführt und die Relevanz des Testgegenstands für die Fahreignung geklärt. Beispielsweise enthält die Handanweisung zum Untertest A eine Definition des Anforderungsbereichs. Dort wird Autofahren als stetiger, rückgekoppelter Informationsfluss verstanden. Theoretische Grundlagen sind auch in den Handanweisungen zum Untertest B, C und D zu finden. Diese werden jedoch nicht in ein theoretisches Modell der Fahreignung, das der gesamten Testbatterie zugrunde liegt, integriert. Eine Zuordnung der Tests bzw. des jeweiligen Testgegenstandes zu einem theoretischen Modell oder zu den Anforderungsbereichen der FeV wird von Seiten der Autoren nicht vorgenommen und fällt somit für Leser oder Anwender schwer.

Zum generellen Aufbau der Manuale des Testsystems A lässt sich sagen, dass eine testspezifische Erläuterung der jeweiligen Relevanz für die Fahr-

eignungsprüfung, eine Abgrenzung der Tests gegeneinander und eine gesonderte Abhandlung der Gütekriterien wünschenswert wären. Häufig finden sich identische Darlegungen im Bezug auf den Anforderungsbereich, dies gilt bspw. für zwei Untertests. Ebenso werden im Manual eines anderen Untertests die Begriffe „Aufmerksamkeit“ und „Konzentration“ synonym verwendet. Im Manual eines anderen Untertests hingegen wird Konzentrationsleistung als „der höchste Grad an Aufmerksamkeitsausprägung“ bezeichnet.

Bei den Gütekriterien wird meist auf Veröffentlichungen im Anhang der Manuale verwiesen. Diese beziehen sich jedoch auf die gesamte Testbatterie A und thematisieren deren Funktionalität, Ergonomie, die Prinzipien der Hardware- und Softwaregestaltung, empirische Überprüfungen der verschiedenen Instruktionversionen und die empirische Überprüfung einer fremdsprachigen Version des Testsystems A. Somit ist das notwendige Informationsmaterial gewissermaßen vorhanden, muss aber vom Leser aufwändig zusammengetragen werden. Dies erschwert eine genaue Analyse der einzelnen Tests und bietet wenig Transparenz in Bezug auf das einzelne Verfahren. Interpretationen oder Äußerungen zu den Gütekriterien von Seiten der Autoren im Manual selbst wären wertvolle Hinweise für die Beurteilung der Tests.

Zur Normierung ist in den Handanweisungen der einzelnen Tests zu lesen, dass sich die Stichprobe für die Normierung „aus dem Klientel der verkehrspsychologischen Fahreignungsdiagnostik“ zusammensetzte. Wie bereits erwähnt, werden die Normstichproben jedoch innerhalb der einzelnen Handanweisungen nicht detailliert beschrieben, sondern auf Veröffentlichungen im Anhang verwiesen. Außerdem werden die Normen selbst nicht im Anhang angeboten. Dadurch wird die Möglichkeit, die Normwerte PR 16 und 33 direkt zu ermitteln, nicht gegeben. Bei einzelnen Tests ist auch ein Hinweis, dass die Normbasis „laufend erweitert“ wird, zu finden. Insgesamt wird Anwendern so kein selbstständiger und kritischer Umgang mit den Normen ermöglicht.

Bei der Durchführung der einzelnen Testanalysen konnte lediglich angenommen werden, dass die einzelnen Normierungsstichproben sich entsprachen. Genaue Angaben fehlten jeweils. Dennoch ist zu vermuten, dass die Zusammensetzung der Eichstichprobe nicht der Grundgesamtheit aller Führerscheininhaber in Deutschland entsprach. Neben

dem Aspekt, dass es sich um eine nicht-deutsche Stichprobe handelt, gehen die Autoren hier nämlich grundsätzlich von einer anderen Zielpopulation aus. Nach der Lektüre aller Manuale des Testsystems A ist anzunehmen, dass sich die Stichprobe nur aus männlichen Klienten zusammensetzte und dass die Personen aus den so bezeichneten „Hauptanlassgruppen Alkoholauffälligkeit, Berufskraftfahrer, Drogenlenker und verkehrsauffällige ältere Kraftfahrer“ stammten. Somit verstehen die Autoren das Klientel der verkehrspsychologischen Fahreignungsdiagnostik als Zielpopulation.

In Bezug auf die Validierung, insbesondere auf die Kriteriumsvalidierung der Tests des Systems A, fiel die Testanalyse ebenfalls schwer. Zwar werden vielerorts Studien zum Zusammenhang mit standardisierten Fahrproben angegeben, doch sind immer wieder unterschiedliche Bezeichnungen unterschiedlicher Studien zu finden. Da diese Bezeichnungen und Studien jedoch an keiner Stelle explizit eingeführt werden, sind sie für Anwender und Leser nicht aussagekräftig. Auch hier wäre eine detaillierte Beschreibung der Validierung bezogen auf den einzelnen Untertest wünschenswert. Des Weiteren wäre auch im Hinblick auf die Validität ein einheitliches theoretisches Modell erstrebenswert.

Durch die fehlende Transparenz bei der Beschreibung der unterschiedlichen Stichproben konnte bei den Testanalysen nicht beurteilt werden, ob die Angaben zur Reliabilität der Tests auf der Stichprobe der Eichung beruhen. Als Reliabilitätsmaß wird häufig Cronbachs Alpha angeführt, allerdings ohne Angabe von zugrunde liegenden Itemzahlen o. Ä.

Basierend auf den definierten objektiven Kriterien würden somit folgende Aspekte die Anwendung des Testsystems A erleichtern:

- Die Darlegung eines theoretischen Modells der Fahraufgabe bzw. der Fahreignung, genaue Verwendung von Definitionen und Bezeichnungen der Anforderungsbereiche und eine entsprechende Herleitung der Testverfahren.
- Eine darauf aufbauende genaue Zuordnung von Tests zu Anforderungsbereichen nach FeV.
- Eine Empfehlung, welche Testversion sich jeweils am besten eignet (bei einem Untertest wird beispielsweise nicht transparent, welches die Standardversion ist, auf welche Version sich die Gütekriterien beziehen, wie es sich mit den Gütekriterien der übrigen Versionen verhält und

welche für die Prüfung der Fahreignung am besten geeignet ist).

- Eine Einschränkung von automatischen Gutachten, da bereits bestehende Textbausteine das differenzierte Auseinandersetzen mit dem Einzelfall durch den Gutachter verhindern können.

Eine schärfere Abgrenzung der Anforderungsbereiche und theoretischen Grundlagen der Untertests des Testsystems A würde dem Anwender somit eher eine gezielte Anwendung einzelner Tests auf besondere Fragestellungen hin und einen kritischen und flexiblen Umgang ermöglichen.

7.3 Testsystem B

Das Testsystem B besteht ebenfalls aus verschiedenen Untertests.

Die nun folgende zusammenfassende Beurteilung des Testsystems B beruht ebenfalls auf den Testanalysen zu den Untertests und dem übergreifenden Handbuch zum Testsystem B.

Aufbauend auf dem Handbuch des Testsystems B selbst lässt sich feststellen, dass dieses Testsystem auf einem Modell des Fahrverhaltens basieren soll. Eine explizite wissenschaftliche Herleitung der Einzeltests wird jedoch nicht durchgeführt oder dargestellt. Das Handbuch enthält Bezüge zu gesetzlichen Bestimmungen in verschiedenen europäischen Ländern. In den Manualen der Einzeltests wird dies jedoch nicht verdeutlicht.

Insgesamt lässt sich das Testsystem B als anwenderfreundlich ausweisen, da es Interpretationshilfen, vor allem in Abstimmung mit den Begutachtungs-Leitlinien der Bundesanstalt für Straßenwesen (BASt), Empfehlungen für konkrete Testversionen, Screenshots der Auswertung und praktische Fallbeispiele aufweist. Allerdings gilt es, auch hier kritisch anzumerken, dass durch die Möglichkeit der Erstellung automatisierter Gutachten bei bereits vorgegebenen Textbausteinen die Gefahr besteht, dass sich Gutachter nicht mehr differenziert mit dem Einzelfall auseinandersetzen.

Verbesserungswürdig wäre auch die Tatsache, dass die in den einzelnen Testmanualen zusätzlich genannten testspezifischen theoretischen Grundlagen nicht in das angenommene Grundmodell des Fahrverhaltens (s. o.) integriert werden. In den Manualen einiger Untertests bleibt der Kontext der Fahreignungsdiagnostik oder verkehrspsychologi-

schen Diagnostik weitgehend unerwähnt. Hilfreich erscheinen hingegen die übergeordnete Zuordnung von Tests des Testsystems B zum zugrunde liegenden Modell der Fahraufgabe einerseits und die Zuordnung der Tests zu den Anforderungsbereichen und ihre Zusammenstellung zu einer Testbatterie nach FeV Anlage 5 andererseits. Somit werden auch fehlende Empfehlungen innerhalb der einzelnen Testmanuale, welcher Tests welchem Funktionsbereich nach FeV entspricht, kompensiert. Hier wird versucht, den Vorgaben des Gesetzgebers entsprechend eine Empfehlung für Anwender zu geben. Allerdings enthält auch das Testsystem B über Aufmerksamkeitsaspekte hinausgehende und somit den Anforderungsbereichen der FeV, Anlage 5, nicht zuzuordnende Tests.

Im Hinblick auf die Normierung wird innerhalb des Testsystems B eine – auf Ergebnissen von Volkszählungen in verschiedenen europäischen Ländern basierende – repräsentative Stichprobe der Normalbevölkerung angegeben. Je nach Test und dort wiederum abhängig von der Testunterform sind die Normen jedoch auch getrennt nach Alter, nach Geschlecht, für Kinder und teilweise auch für verkehrsauffällige Personen verfügbar. Diese stammen zum Teil ebenfalls aus unterschiedlichen Ländern und unterschiedlichen Einrichtungen. Somit fällt ein Überblick über die Normierungslage des Testsystems B streckenweise schwer.

Bei den Tests, bei denen für die Normierung repräsentative Stichproben der Normalbevölkerung angegeben werden, kann davon ausgegangen werden, dass sie weitgehend der Grundgesamtheit aller Führerscheininhaber in den genannten europäischen Ländern entsprechen. Dies ist im Unterschied zu anderen Testsystemen zu sehen. Da es sich beim Testsystem B jedoch um eine Zusammenstellung von Tests verschiedener Testautoren handelt, sind hier Unterschiede zu finden. Beispielsweise beruht die Normierung eines Untertests auf Stichproben aus drei verschiedenen Ländern mit einer Altersspanne von 15 bis 77 Lebensjahren.

Zur Validierung und insbesondere Kriteriumsvalidierung werden häufig Studien zum Zusammenhang mit standardisierten Fahrproben oder Extremgruppenvergleiche angeführt und beschrieben. Auch hier sind Unterschiede aufgrund von unterschiedlichen Testautoren zu finden.

Im Hinblick auf die Reliabilität wird häufig Cronbachs Alpha angeführt. Meist wird jedoch nicht diskutiert, wie mit Auslassern verfahren wurde oder

welche Itemzahl der tatsächlichen Berechnung zugrunde liegt. Auch basieren nicht bei allen Tests die Reliabilitätsmaße auf den Eichstichproben.

Zusammenfassend würden auf Basis der formulierten objektiven Kriterien folgende Aspekte die Anwendung des Testsystems B erleichtern:

- Eine Darstellung, auf welche Weise die Tests tatsächlich wissenschaftlich hergeleitet wurden und eine Überführung der angegebenen theoretischen Grundlagen in die einzelnen Testmanuale.
- Eine übersichtliche Darlegung der Normierung der einzelnen Tests, der dem Einzeltest zugrunde liegenden Eichstichproben, insbesondere in Bezug auf die jeweiligen für die Fahreignungsprüfung relevanten Testformen.
- Eine Einschränkung von automatischen Gutachten, da die bereits vorgegebenen Textbausteine das differenzierte Auseinandersetzen mit dem Einzelfall durch den Gutachter verhindern können.

7.4 Testsystem C

Das Testsystem C besteht ebenfalls aus mehreren Untertests.

Beim Testsystem C gilt es zunächst festzuhalten, dass theoretische Grundlagen zur Aufmerksamkeitsforschung in Verbindung mit der Fahreignung und eine kritische Auseinandersetzung mit den Anforderungsbereichen FeV, Anlage 5, dargestellt werden. Leider wird dies jedoch nicht im Aufbau des Manuals fortgeführt. Dieser könnte konsequenter dem Ziel der Fahreignungsprüfung folgen. Anfänglich wird eine Zuordnung der theoretischen Grundlagen zu Anforderungsbereichen der FeV vorgenommen. Anschließend ist aber bei den verschiedenen Tests kein Bezug zur vorangegangenen Zuordnung oder einem eigenen Modell der für die Fahreignung zentralen Aufmerksamkeitsaspekte zu finden. Dabei werden Defizite einzelner kognitiver Funktionen detailliert besprochen, jedoch ohne dass deren Relevanz für die Fahrtauglichkeit explizit erläutert wird. Empfehlenswert wäre ein Aufbau des Manuals getrennt nach Untertests, die für die Überprüfung der Fahreignung nach FeV empfohlen werden, und den übrigen Tests, deren Relevanz ebenfalls erläutert wird.

Im Testsystem C wird eine Empfehlung der Zuordnung von Untertests zur Anwendung im Hinblick auf

die Anforderungsbereiche nach FeV, Anlage 5, vorgenommen. Dabei wird im Manual kritisch angemerkt, dass sich diese Anforderungsbereiche „nur begrenzt mit aktuellen Taxonomien der Aufmerksamkeit verknüpfen“ lassen.

Insgesamt scheint Testsystem C die Fahreignung ausgehend von medizinischen Risikogruppen zu betrachten, da der Fokus der Anwendung auf Fahreignungsuntersuchung im Kontext von verschiedenen Krankheiten liegt.

Zur Normierung lässt sich ein abschließendes Urteil nicht fällen. Zum einen ist die Normierung der „neueren Tests“ des Testsystems C noch nicht erfolgt. Zum anderen gibt es keine Jahresangabe zu den Normierungen der alten Untertests. Auch die Frage, ob die Probanden der Eichstichproben überhaupt im Besitz eines Führerscheins waren oder nicht, geht aus den Manualen nicht hervor. Allerdings wird vom Testsystem C die Notwendigkeit von altersunabhängigen (nicht alterskorrigierten) Normen angesprochen.

Den angegebenen Studien zur Kriteriumsvalidität liegen häufig spezielle Probandengruppen zugrunde und es wird meist nicht das unmittelbare Fahrverhalten als Kriterium genommen.

Da die Normierung der „neu entwickelten Tests“ noch nicht erfolgt ist, liegen noch keine Reliabilitätsmaße für diese Tests vor. Die Maße, die in Bezug auf die übrigen Tests aufgeführt werden, werden kritisch betrachtet: „Während Split-Half relativ sensitiv bezüglich interindividuellen Trendunterschieden ist, kann Odd-Even durch Ausreißer beeinflusst werden. Daher wurden für die vorhandenen Tests beide Koeffizienten berechnet.“

Nachstehende Aspekte würden demnach die Anwendung des Testsystems C erleichtern:

- Ein am Ziel der Fahreignungsüberprüfung nach FeV ausgerichteter Aufbau des Manuals.
- Die Vervollständigung der Normierungs- und Reliabilitätsstudien.
- Normierung entsprechend der Grundgesamtheit der Gruppe der Führerscheinbesitzer, die mit dem Testsystem gemessen werden soll.

7.5 Testsystem D

Das Testsystem D umfasst ebenfalls verschiedene Einzeltests.

In Bezug auf das Testsystem D ist in erster Linie die mangelnde Verfügbarkeit von Informationen festzustellen. Die Testanalyse des Testsystems D basiert auf einer „Broschüre“ zum Testsystem und der zum Testsystem gehörenden „Experten-Information“, wobei sich Letztere kürzer fasst, einen anderen Aufbau aufweist, jedoch keine zusätzlichen Informationen beinhaltet. Die Unterlagen erwiesen sich als unzureichend, da wesentliche Elemente und Testgütekriterien, die das Manual eines wissenschaftlichen Testes enthalten sollte, fehlen.

Generell liegt der Fokus der Unterlagen zum Testsystem D auf seiner Darstellung in Abgrenzung von anderen, nicht näher benannten Verfahren der Fahreignungsdiagnostik. Diese allein ist aus wissenschaftlicher Perspektive jedoch eher unwesentlich. Bezogen auf das eigentliche Testsystem D werden keine oder sehr vereinfachte theoretische Grundlagen in Bezug auf die Fahraufgabe dargestellt und es fehlt eine Einbettung der Relevanz der aufgeführten Aufmerksamkeitsfunktionen in die Fahraufgabe. Es wird vielmehr explizit eine Augenscheininvalidität der Testverfahren angenommen, die ohne weitere wissenschaftliche Herleitung als ausreichend betrachtet wird.

Eine Zuordnung der Anforderungsbereiche nach FeV zu Funktionen, die mit Untertests des Testsystems D erfasst werden sollen, wird vorgenommen, jedoch nicht detailliert dargelegt oder begründet. Und auch hier dient der Vergleich mit anderen Verfahren als Richtschnur. In Bezug auf den Anforderungsbereich „Orientierungsfähigkeit“ lautet z. B. die Aussage, dass es sich vertreten lasse, das Merkmal nur indirekt zu erfassen, weil dieses Merkmal von anderen Testsystemen ebenfalls nur indirekt erfasst werde.

Als Normdatensätze werden zwei Datensätze genannt: ein Datensatz einer hochspezialisierten Berufsgruppe und ein weiterer Datensatz, der im Rahmen universitärer Studienabschlussarbeiten gewonnen wurde. Der Erhebungszeitraum der Normierung der hochspezialisierten Berufsgruppe lag im Jahr 1996. Bezüglich der Datenerhebung für die universitäre Studienabschlussarbeit wird keine Jahresangabe gemacht.

Mit den angegebenen beiden Normierungsstichproben wird nicht die Grundgesamtheit aller Führerscheininhaber in Deutschland als Zielpopulation angenommen. Wiederum nimmt der Autor eine Abgrenzung gegen verbreitete Vorstellungen vor, dass Normstichproben möglichst groß sein müssten,

ohne die eigene vorliegende Normierung zu begründen oder näher zu erläutern.

Der Autor der Broschüre zum Testsystem D kritisiert die besondere Berücksichtigung der Validierung von Testverfahren am Fahrverhalten und die häufig genannten geringen Zusammenhänge bzw. Korrelationskoeffizienten. Als Erklärung liefert er statistische Aspekte, wie die Größe der Stichprobe, die die Häufigkeit von Signifikanzen beeinflusst, die homogene Konstruktion von Tests im Hinblick auf deren Reliabilität und die unterschiedlichen Verhaltensdispositionen, die dem Testverhalten zum einen und dem Fahrverhalten zum anderen zugrunde liegen. Er betont die Notwendigkeit der internen Validierung. Trotz dieser Einleitung führt der Autor anschließend einen Nachweis anhand eines Extremgruppenvergleichs an. Hierbei handelt es sich um Unterschiede in den Testleistungen von Personenbeförderern und einem in der Altersverteilung vergleichbaren Teil der Normstichprobe. Eine Validierung anhand von standardisierten Fahrproben wird demnach nicht geliefert.

Als Reliabilitätsmaße werden Kuder/Richardson-Alpha und Split-Half-Reliabilitäten aufgeführt. Die zugrunde liegende Stichprobe wird jedoch nicht beschrieben und die Auswahl der Reliabilitätsmaße nicht begründet oder dargelegt.

Demnach wären auf Basis der formulierten objektiven Kriterien folgende Aspekte für die Anwendung des Testsystems D im Rahmen der Fahreignungsprüfung von grundlegender Bedeutung:

- Die Gestaltung eines für Verfahren der wissenschaftlichen Diagnostik üblichen Manuals.
- Eine wissenschaftliche Fundierung und Herleitung des Testsystems und der Untertests.
- Die Darstellung der wissenschaftlich üblichen Testgütekriterien.
- Die Normierung der Einzeltests an einer für Autofahrer repräsentativen Stichprobe mit einem entsprechenden Umfang.
- Die Darlegung eines theoretischen Modells der Fahraufgabe bzw. der Fahreignung, klare Verwendung von Definitionen und Bezeichnungen der Anforderungsbereiche.

7.6 Test E

Test E ist ein Einzeltest. Der Test E wird als allgemeiner Leistungstest verstanden, der allgemeine

Voraussetzungen zur Erzielung von Leistung erfasst. Mit allgemeinen Voraussetzungen sind Sachverhalte wie Konzentration, Aufmerksamkeit und Aktivierung gemeint. Auch in der letzten überarbeiteten und neu normierten Auflage des Test E wird der zuvor geäußerten wissenschaftlichen Kritik an Details Rechnung getragen und entsprechend ausführlich besprochen. Studien zur Reliabilität und Normen werden ebenfalls geliefert.

Der Test E wurde ursprünglich für Zwecke der Fahreignungsprüfung konzipiert. Dennoch fehlen theoretische Grundlagen zur Fahraufgabe oder Bezüge zu den Anforderungsbereichen der FeV. Beispielsweise fallen hier die Anforderungsbereiche Konzentrationsleistung und Aufmerksamkeit zusammen. Die Tätigkeiten im Test sollen ein hohes Maß an Aufmerksamkeit und Konzentration, aber keine speziellen Fähigkeiten und Fertigkeiten erfordern.

Die neueste deutsche Normierung des Tests E beruht auf einer Stichprobe aus dem Zeitraum 11/1999 bis 05/2000. Des Weiteren liegen Normen basierend auf einer großen Stichprobe der Bundesanstalt für Arbeit sowie französische und berufsgruppenspezifische Normen vor. Bei der Zusammensetzung der Eichstichprobe fällt an dieser Stelle allerdings eine Wertung schwer, da nicht klar wird, wie die „zufallsgesteuerte Stichprobentechnik“ und die Erhebung „vorwiegend an Schulen, in Betrieben, Psychologischen Universitätsinstituten und anderen öffentlichen Einrichtungen“ zu vereinbaren sind. Führerscheinbesitz scheint demnach kein Kriterium gewesen zu sein, sodass davon ausgegangen werden kann, dass die Normierungsstichprobe nicht der Grundgesamtheit aller Führerscheininhaber in Deutschland entsprach.

In Bezug auf die Validierung des Tests E ist festzustellen, dass sich der Gebrauch von Validitätsbegriffen im Test E-Manual von der hier zugrunde gelegten Verwendungsweise unterscheidet. Das Manual unterscheidet Psychologische Validität, Konstruktvalidität und Empirische Validität. Unter Empirischer Validität werden die Übereinstimmungsvalidität und die Vorhersagevalidität gefasst. Diese entsprechen weitgehend der hier interessierenden Kriteriumsvalidität. Jedoch werden keine Studien zum Zusammenhang der Leistung in Fahrproben und der Leistung bei Test E aufgeführt. Es wird lediglich erwähnt, dass der Test E im Rahmen der Überprüfung der Fahreignung eingesetzt wird. Außerdem ist hier festzuhalten, dass der überwiegende Teil der Validitätskriterien nicht ebenfalls im Zuge der neunten Auflage überarbeitet wurde.

Es werden Angaben zur Reliabilität gemacht, die sich sowohl auf ältere Daten als auch auf die neue Eichstichprobe beziehen. Reliabilitäten werden auch für einzelne Altersgruppen, die für die Normierung herangezogen wurden, angegeben. Berücksichtigt werden Cronbachs Alpha und die Split-Half-Reliabilität (Spearman-Brown-korrigiert).

In Bezug auf die Nützlichkeit des Tests E für die Fahreignungsprüfung wären folgende Ergänzungen wünschenswert:

- Die Darlegung eines theoretischen Modells der Fahraufgabe bzw. der Fahreignung und eine Zuordnung des Tests zu den Anforderungsbereichen der FeV.
- Die Nachlieferung von Validierungsstudien, insbesondere zur Kriteriumsvalidität anhand von Fahrverhaltensbeobachtung.

7.7 Test F

Test F ist ebenfalls ein Einzeltest und vor allem für die Anwendung in der Schulpsychologie konzipiert. Deshalb ist keine Beschreibung seiner Relevanz für die Verkehrspsychologie zu finden. Wie beim Test E fallen auch hier die Anforderungsbereiche Konzentrationsleistung und Aufmerksamkeit zusammen.

Die Normierung wird vor allem für Schüler vorgenommen. Es werden keine näheren Angaben zur Stichprobenwahl für die Erwachsenen-Normen gemacht. Somit ist die angenommene Zielpopulation nicht die Gesamtheit aller Führerscheininhaber in Deutschland.

Bezüglich der Validierung sind lediglich Studien zur Kriteriumsvalidierung anhand von Lehrerurteilen oder Schulnoten zu finden.

Auch die Reliabilitätsmaße werden für verschiedene Schulklassen aufgeführt. Es werden Angaben zur Reliabilität früherer Versionen des Tests F, der verschiedenen Maße einer abgewandelten Form des Tests, der beiden Schwierigkeitsformen und der jeweiligen Parallelformen gemacht. Es fehlen jedoch Angaben zu Erwachsenenstichproben.

Beim Test F gilt es schließlich festzustellen, dass er trotz seiner ursprünglichen Ausrichtung keinen Bezug zur Fahraufgabe aufweist und somit keine Relevanz für die umfassende Überprüfung der Fahreignung nach FeV hat.

7.8 Interpretation der Anforderungsbereiche durch verschiedene Testautoren in der Praxis

Wie die einzelnen Testanalysen und die zusammenfassenden Darlegungen zu den Testverfahren in diesem Kapitel zeigen, weist die gegenwärtige Fassung der FeV, Anlage 5, Interpretationsspielraum für Anwender und Testentwickler auf. Häufig werden zwei Anforderungsbereiche synonym verwendet. Dies gilt bspw. für die Anforderungsbereiche Konzentrationsleistung und Aufmerksamkeit im Testsystem A.

Darüber hinaus werden im Testsystem A die Anforderungsbereiche der FeV, Anlage 5, nicht thematisiert. Die angebotenen Tests sind nicht genau deckungsgleich mit den Anforderungsbereichen, sie sind bei einigen Untertests einerseits redundant und bei anderen Untertests andererseits weitergefasst als die Aufmerksamkeitsaspekte der FeV.

Das Testsystem B hat einen breiteren Blickwinkel und bündelt die Anforderungsbereiche zumindest vordergründig in ein Modell der Fahraufgabe ein. Die Interpretation der Anforderungsbereiche der FeV und die Festlegung von Tests zu deren Überprüfung können dem Manual des Testsystems entnommen werden.

Testsystem C würdigt ebenfalls die Anforderungsbereiche der FeV, aber distanziert sich zugleich. Eine Zuordnung von Testgegenständen zu Anforderungsbereichen wird vorgenommen, gleichzeitig wird jedoch auch die mangelnde Definitionsschärfe der Anforderungsbereiche angemerkt.

Auch im Testsystem D wird eine Zuordnung vorgenommen. Dort heißt es in Bezug auf die Anforderungsbereiche „...das Merkmal 'Aufmerksamkeit' ist darin explizit enthalten, andere lassen sich mit bestimmten, durch Testsystem D erfassten Funktionen gleichsetzen bzw. als eine Auslegung der gemeinten Fähigkeit auffassen.“

Bei den beiden Einzeltestverfahren Test E und Test F wurde bereits festgestellt, dass Anforderungsbereiche wie Konzentrationsleistung und Aufmerksamkeit zusammenfallen.

Schlussfolgernd ist hier also festzustellen, dass die aktuelle Fassung der FeV, Anlage 5, Testautoren zu unterschiedlichen Interpretationen führt und dass es in Zukunft einer genaueren Definition und Eingrenzung der für die Fahreignung zentralen Anforderungsbereiche bedarf.

derungen bedarf. Es gilt folglich, den Interpretationsspielraum zu verringern und die Fahreignungsprüfung weitergehend zu standardisieren.

7.9 Einteilung der Tests in drei Kategorien

Der unter Kapitel 4.4 vorgesehene Arbeitsschritt zur Einteilung der Tests in drei Kategorien kann abschließend wie folgt vorgenommen werden: Keiner der Tests bzw. der Testsysteme genügte den Kriterien der Gruppe 3 (Tests, die den in der FeV genannten Kriterien entsprechen). Somit verteilen sich die betrachteten Verfahren auf die Gruppen 1 und 2.

Wie bereits beschrieben, sind alle Tests bzw. Testsysteme unzureichend normiert, da sie meist nicht die Grundgesamtheit aller Führerscheininhaber in Deutschland als Zielpopulation repräsentieren.

Die Tests des Testsystems B weisen noch am ehesten einige Kriterien der Gruppe 2 auf: Sie entsprechen zum Teil thematisch den in der FeV genannten Bereichen und sind darüber hinaus meist auch entsprechend normiert. Sie weisen jedoch nicht durchgehend ausreichende Validierungsstudien auf und auch die Normierung der Untertests unterscheidet sich und muss je nach empfohlener Testform gesondert herausgelesen werden.

Im Wesentlichen kann jedoch gesagt werden, dass alle betrachteten Verfahren – mit Ausnahme des Tests E – eher der Gruppe 1, d. h. Tests, die thematisch den in der FeV genannten Bereichen entsprechen, jedoch nicht entsprechend normiert wurden und für die keine ausreichenden Validierungsstudien vorliegen, zuzuordnen sind.

7.10 Fazit der Testanalysen

Die Testanalysen haben insgesamt gezeigt, dass bei allen analysierten Testsystemen in unterschiedlichem Ausmaß methodische Schwächen bestehen. Diese beziehen sich auf Aspekte der Normierung, der Objektivität, der Reliabilität, der Validität sowie auf die Dokumentation, die häufig sehr lückenhaft ist. Darüber hinaus fehlen häufig Angaben zur Testentwicklung.

Abschließend gilt es festzuhalten, dass eine Schärfung der FeV im Hinblick auf die Prüfung der Fahreignung und der zu diesem Zweck eingesetzten Verfahren in den eingangs beschriebenen Aspekten notwendig ist. Die Testanalysen haben gezeigt,

dass die Manuale selten einem Standardaufbau folgen. Dies erschwert die kritische Betrachtung und Anwendung von Tests. So wäre neben der Schärfung der Definition der Anforderungsbereiche selbst auch eine Konkretisierung der Anforderung an Testverfahren zu fordern. Die Begriffe der „wissenschaftlichen Standardisierung“ und „Validierung unter Aspekten der Verkehrssicherheit“ gilt es weiter auszuarbeiten und Vorgaben für die Gestaltung von Manualen zu machen. Beispielsweise sollte die Darlegung von theoretischen Grundlagen und der Relevanz für die Prüfung der Fahreignung darin enthalten sein.

8 Expertentreffen: Protokoll des Expertengesprächs zu den Anforderungen der Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung

Am 21. März 2007 fand im Institut für Arbeitsphysiologie an der Universität Dortmund (IfADo) ein Expertentreffen statt, bei dem der damalige Projektstand (das ist im Wesentlichen der in den vorangegangenen Kapiteln dargestellte Erkenntnisstand) einen Tag lang mit von der BAST eingeladenen Experten aus Forschung und Praxis besprochen wurde.

Das Protokoll wurde als Ergebnisprotokoll mit einer strukturierten Zusammenstellung der wesentlichen Gesprächsinhalte und zum Teil namentlicher Zuordnung erstellt. Für die Teilnehmer bestand die Möglichkeit, Änderungswünsche in das Protokoll einzubringen. Hiervon wurde nur in einem Fall in geringem Umfang Gebrauch gemacht.

8.1 Teilnehmerübersicht

Jürgen Brenner-Hartmann,
Leiter des Arbeitskreises „Beurteilungskriterien“

Prof. Dr. Michael Falkenstein,
Projektgruppe „Altern und ZNS-Veränderungen“,
Institut für Arbeitsphysiologie an der Universität
Dortmund

Dr. Frank Goldhammer,
Institut für Psychologie, Universität Frankfurt

Dr. Hendrik Niemann,
NRZ Leipzig, Neuropsychologie

Dr. Sebastian Poschadel,
Projektgruppe „Altern und ZNS-Veränderungen“,
Institut für Arbeitsphysiologie an der Universität
Dortmund

Prof. Dr. Georg Rudinger,
Methodenlehre und Diagnostik, Psychologisches
Institut der Universität Bonn

Andreas Schale,
Abteilungsleitung Psychologie, Dr. Becker Klinikge-
sellschaft mbh & Co. KG

Prof. Dr. Lothar Schmidt-Atzert,
Differenzielle Psychologie und Psychologische Dia-
gnostik, Universität Marburg

Prof. Dr. Edmund Wascher,
Kognitive Ergonomie, Institut für Arbeitsphysiologie
an der Universität Dortmund

Prof. Dr. Klaus Willmes-von Hinckeldey,
Lehr- und Forschungsgebiet Neuropsychologie an
der Neurologischen Klinik, RWTH Aachen

BAST

Dr. Anja Knoche,
Referat „Verkehrspsychologie, Verkehrsmedizin“

Dr. Horst Schulze,
Leiter des Referats „Verkehrspsychologie, Ver-
kehrsmedizin“

Moderation

Hardy Holte (BAST),
Referat „Verkehrspsychologie, Verkehrsmedizin“

8.2 Begrüßung und Vorstellungsrunde

Professor Falkenstein begrüßt die Teilnehmer zur Gesprächsrunde im Institut für Arbeitsphysiologie an der Universität Dortmund und weist darauf hin, dass Herr Holte von der Bundesanstalt für Straßenwesen die Moderation dieser Gesprächsrunde übernehmen wird. Er weist außerdem darauf hin, dass Herr Dr. Poschadel im Anschluss an eine kurzen Problemdarstellung von Herrn Holte vorläufige Projekt-Zwischenergebnisse vortragen wird. Ziel dieser Gesprächsrunde ist es, diese Ergebnisse

ausführlich zu diskutieren und die Diskussionsresultate in das Projekt einfließen zu lassen. Professor Falkenstein bittet die Teilnehmer, sich kurz vorzustellen.

8.3 Problemstellung

Vorgetragen von Herrn Holte

8.3.1 Vorgeschichte des Projekts

Im Rahmen der Fahreignungsdiagnostik wird eine Reihe von Testsystemen angewendet, um bestimmte Anforderungsbereiche zu testen, die in der Anlage 5 (2) der Fahrerlaubnisverordnung (FeV) zur Erfassung der psychischen Leistungsfähigkeit vorgegeben sind (Belastbarkeit, Orientierungsleistung, Konzentrationsleistung, Aufmerksamkeitsleistung, Reaktionsfähigkeit).

Im Jahr 2000 hat es in der BAST ein Expertengespräch gegeben, bei dem es um die Klärung der Frage ging, welche wissenschaftlich-methodischen Anforderungen an die Entwicklung von Tests zu stellen sind, die in der Fahreignungsdiagnostik auf der Grundlage der Anlage 5 (2) FeV angewendet werden. Ein zentrales Ergebnis dieses Expertentreffens war der Vorschlag, ein Projekt zur Klärung der offenen Fragen durchzuführen. Darüber hinaus erhielt die BAST eine Reihe von Anfragen aus den Ländern, mit der Bitte, alle Tests aufzulisten, die in der Fahreignungsdiagnostik aktuell angewendet werden.

In einem Erlass vom BMVBW (Bundesministerium für Verkehr, Bau- und Wohnungswesen) im Jahr 2001 wurde die BAST schließlich gebeten, zu dieser Thematik eine Projektbeschreibung zu erstellen. Im Jahr 2005 wurde dieses Projekt in das Sicherheitsprogramm aufgenommen und im Rahmen eines üblichen Vergabeverfahrens der BAST im Dezember 2005 an das Institut für Arbeitsphysiologie an der Universität Dortmund (IfADo) vergeben.

8.3.2 Zielsetzung des Projekts

Der Fokus in diesem Projekt liegt auf einer Analyse der in der Fahreignungsdiagnostik verwendeten Testsysteme anhand der in der psychologischen Diagnostik verwendeten Testgütekriterien. Vier zentrale Arbeitsschritte sind vorgesehen:

1. Bestandsaufnahme: Testrecherche (Welche Tests werden eingesetzt?), Befragung von

MPU-Gutachtern bezüglich der Anwendung der Tests in der Praxis (z. B.: Welche Tests werden in der jeweiligen Begutachtungsstelle für Fahreignung verwendet?),

2. Testanalyse und -bewertung: Durchsicht der Testmanuale und der Literatur (Gütekriterien),
3. Expertengespräch zur Diskussion der Ergebnisse aus Schritt 1 und 2,
4. Vorschläge für Weiter- und Neuentwicklung von standardisierten Tests zur Begutachtung der Fahreignung. Empfehlungen zur Anwendung der Begutachtungsleitlinien unter methodisch-neuropsychologischen Gesichtspunkten.

8.4 Präsentation der Projekt-Ergebnisse

Vorgestellt von Dr. Poschadel

Die von Herrn Poschadel vorgetragene Präsentation stellte die Arbeitsgrundlage für das Expertengespräch dar. In der Präsentation wurde der aktuelle und nicht endgültige Bearbeitungsstand des Projektes wiedergegeben.

Zusammenfassend war festzustellen, dass die Zuordnung der Einzeltests durch die Testautoren zu den Konstrukten der FeV nicht immer nachvollziehbar ist. Teilweise wird keine explizite Zuordnung vorgenommen. Dies kann allerdings auch an den unscharfen Konstrukten der FeV, Anlage 5, selbst liegen. Grundsätzlich sollte nachvollziehbar sein (und auch von den Testautoren erklärt werden), was ein Test tatsächlich misst.

Bei den von den Testautoren genannten Reliabilitätsmaßen lässt sich aufgrund der Testmanuale nicht hinreichend klären, welche Stichprobe den Berechnungen zugrunde liegt und aus welchen Gründen welches Reliabilitätsmaß gewählt wurde. Werden Reliabilitäten nach Cronbachs Alpha berechnet, wird nicht angegeben, wie mit Auslassern verfahren wurde.

Insgesamt bestehen in unterschiedlichem Ausmaß methodische Schwächen bei allen analysierten Testsystemen. Diese beziehen sich auf Aspekte der Normierung, der Objektivität, der Reliabilität, der Validität sowie auf die Dokumentation, die häufig sehr lückenhaft ist. Dabei sind auch häufig fehlende Angaben zur Testentwicklung festzustellen.

8.5 Diskussion der Ergebnisse

Herr Holte bittet die Teilnehmer zunächst, die Projektergebnisse und die damit verbundene Vorgehensweise zu kommentieren, und stellt hierzu folgende Eingangsfragen: Welche Leistungsanforderungen sind zu definieren? Wie kommt man zu „guten“ Tests? Was lässt sich zu den Gütekriterien der vorgestellten Tests sagen? Welche Kriterien sollten Validitätsstudien erfüllen? Gibt es ein perfektes Modell der Informationsverarbeitung? Sind die für das Fahren relevanten wissenschaftlichen Konstrukte und Modelle bekannt? Inwieweit korrelieren sie mit dem tatsächlichen Verhalten? Kann man Cut-offs festlegen? Im zweiten Teil der Diskussion liegt der Schwerpunkt auf der vertiefenden Betrachtung der testtheoretischen Kriterien Normierung, Validität/Validierung, Reliabilität und Objektivität.

Die Teilnehmer stimmen darin überein, dass die derzeit in der Fahreignungsdiagnostik eingesetzten psychometrischen Testsysteme aus wissenschaftlicher Perspektive eine Reihe methodischer Schwächen aufweisen. Nachfolgend werden die jeweiligen Diskussionsfelder ausführlicher dargestellt:

1) Theoretische Fundierung der Testentwicklung

Die Teilnehmer stimmten darin überein, dass eine fundierte theoretische Grundlage für die Entwicklung von psychometrischen Tests zur Überprüfung der Fahreignung dringend erforderlich sei. Die Beziehung zwischen Test und wissenschaftlicher Theorie, so Prof. Falkenstein, ist bei den untersuchten Testverfahren oft nicht nachvollziehbar, zum Teil wird eine solche Beziehung nicht explizit hergestellt. Eine unzureichende theoretische Basis hat unmittelbare Auswirkungen auf die für die Fahreignungsdiagnostik definierten Leistungsanforderungen. Bislang sind im Rahmen einer Fahreignungsdiagnose diejenigen Leistungsanforderungen zugrunde gelegt, die in der FeV, Anlage 5 (2), aufgeführt sind: Belastbarkeit, Orientierungsleistung, Konzentrationsleistung, Aufmerksamkeitsleistung, Reaktionsfähigkeit. Übereinstimmend besteht unter den Experten die Ansicht, es müsse ein „neuer Kanon“ mit aktuellen wissenschaftlichen Konstrukten erarbeitet werden, der anschließend in die FeV, Anlage 5, aufgenommen werden kann. Ein solcher Kanon sollte zunächst „Aufmerksamkeit plus“ genannt werden. Das heißt, es sollte zunächst vom kognitiven wissenschaftlichen Modell der Aufmerksam-

samkeit ausgegangen werden, in das weitere für die Fahreignung relevante Konstrukte integriert werden können. Herr Brenner-Hartmann, Prof. Falkenstein und Prof. Rudinger weisen darauf hin, dass fahrrelevante Fähigkeiten nicht auf das Konstrukt „Aufmerksamkeit“ beschränkt werden sollten. So spielen zum Beispiel bestimmte Aspekte der Wahrnehmung und der Motorik eine wichtige Rolle beim Autofahren. Dennoch, so Dr. Goldhammer, könnten die meisten der für die Fahreignung nötigen Fähigkeiten mit dem Konzept „Aufmerksamkeit“ abgedeckt werden. Dazu sei jedoch eine bessere Definition der FeV-Begriffe erforderlich. So stellt sich z. B. die Frage: Was ist Orientierung, was Konzentration? Ist Konzentration mit fokussierter Aufmerksamkeit gleichzusetzen? Dr. Knoche hält das Aufmerksamkeitskonzept zur Erfassung der Fahreignung für besser geeignet als die Konzepte der FeV, es ist jedoch noch nicht klar, ob hierdurch ALLE wesentlichen Teile der Fahreignung abgedeckt werden können. Das Konzept „Aufmerksamkeit“ sollte daher zunächst erweiterbar bleiben. Deshalb müsse man sich bei der Erstellung eines neuen Kanons von Tests nicht unbedingt an der FeV orientieren.

Einige Teilnehmer beklagen den zu großen Interpretationsspielraum für Testentwickler aufgrund der vagen und zum Teil wissenschaftlich überholten Anforderungs-Begriffe (Belastbarkeit etc.), die in der FeV, Anlage 5, genannt werden. Die in der FeV aufgeführten Begriffe würden nicht mit derzeitigen wissenschaftlichen Konstrukten und kognitiven Modellen übereinstimmen, sodass die Testentwickler wahrscheinlich auch wegen dieser schlechten Konstruktlage große Probleme haben, eine Konstruktvalidität anzugeben. Prof. Rudinger hält es für sinnvoll, im Rahmen der Aufstellung eines neuen Kanons von verkehrssicherheitsrelevanten kognitiven Funktionen, eine Expertengruppe zu organisieren.

2) Leistungskriterien

Die Beurteilung der Fahreignung auf Basis der Fahreignungstests erfolgt unter Berücksichtigung der Prozenträge 16 (für alle) und 33 (Gruppe 2, z. B. Bus- und Taxifahrer). Diese Prozenträge sind Setzungen und stellen keine „echten Kriterien“ im Sinne absoluter Kennwerte dar, wie dies zum Beispiel bei der Promilleangabe zum Alkoholgehalt im Blut der Fall ist. Prozenträge basieren auf der Verteilung der Leistungen innerhalb einer Normstichprobe. Ändert sich das Klientel oder die Altersverteilung in der Gesellschaft (demografischer Wan-

del), ändern sich auch die Prozenträge. Die Teilnehmer sprechen sich einstimmig für eine grundsätzliche Hinwendung zu absoluten Kennwerten als Kriterien für bestimmte Leistungsanforderungen aus und fordern damit einen Paradigmenwechsel in der Fahreignungsdiagnostik. Dies ist ein langfristiger Prozess, da die genauen Zusammenhänge zwischen bestimmten kognitiven Funktionen und der Fahreignung noch nicht ausreichend erforscht sind. Dr. Schulze merkt an, dass für Forschungsprojekte, die thematisch darauf ausgerichtet sind, zukünftig gute Chancen bestehen, gefördert zu werden, insbesondere vor dem Hintergrund der größer werdenden Zahl älterer Fahrer.

3) Kompensation von Leistungsdefiziten

In der Diskussion über die für die Fahreignung wichtigsten kognitiven Funktionen wurde auch die Frage nach der Kompensation gestellt. Inwieweit schlechte Leistungen in einem Subtest durch gute Leistungen in einem anderen ausgeglichen werden können, hängt von dem jeweiligen Gewicht der einzelnen Funktionen (Subtests) ab. Eine solche Gewichtung – und damit die Verrechnung verschiedener Scores – ist derzeit weder theoretisch ableitbar noch empirisch hinreichend belegt. Von den Teilnehmern unbeantwortet blieb die Frage von Prof. Willmes, wie mehrere knapp überschwellige Subtest-Ergebnisse mit einem stark unterschwelligen Ergebnis verrechnet werden können.

Darüber, dass bestimmte schlechte Leistungen in einem Subtest eines Testsystems durch gute Leistungen in einem anderen Subtest ausgeglichen werden können, besteht unter den Teilnehmern kein Zweifel. Zum Teil unklar ist jedoch, ob die in der Praxis diagnostizierten Kompensationsmöglichkeiten tatsächlich zum Erhalt der Fahreignung beitragen. Einen ausreichenden empirischen Beleg hierfür scheint es bislang nicht zu geben. Dr. Niemann weist darauf hin, dass bei Patienten andere Kompensationsstrategien bestehen als bei Gesunden. Die Frage, wie man feststellen kann, welche defizitären Leistungen durch welche verkehrssicherheitsrelevanten Leistungen in anderen kognitiven Bereichen kompensierbar sind, lässt sich nur empirisch beantworten. In diesem Zusammenhang weist Herr Holte auf die Möglichkeit einer Profilvalidierung (s. auch Punkt 5 „Validität und Validierung“) hin, bei der nicht ein einziger, spezifischer Test auf Validität geprüft wird, sondern Ergebnisprofile, die im Zusammenhang mit Fahrverhalten oder Unfällen stehen.

4) Umfang und Zusammensetzung von Normstichproben

In den Begutachtungsleitlinien wird von altersunabhängigen Normwerten gesprochen. Altersunabhängig, so wurde von den Teilnehmern bestätigt, ist als repräsentativ zu verstehen. Dieser Anspruch wird gegenwärtig nur zum Teil erfüllt. Von den derzeitigen Testverfahren liegen häufig nur Daten von Auffälligen, von Auffälligen gemischt mit Normalen oder von Extremgruppen vor. Die Frage, die immer wieder gestellt wird, lautet: Darf man das Klientel der Auffälligen als Eichstichprobe heranziehen? Es besteht Konsens zwischen den Teilnehmern, dass die Eichstichprobe repräsentativ für die Grundgesamtheit der Population sein muss, für die sie gelten soll (also für den potenziellen Kraftfahrer). Es genügt nicht, das Klientel der Auffälligen als Eichstichprobe heranzuziehen. Allerdings müssen für die Testung von Gruppen mit besonderen Anforderungen (Gruppe 2, z. B. Busfahrer) Normstichproben aus eben dieser Gruppe gebildet werden.

Die Teilnehmer sprechen sich gegen die Einführung altersspezifischer Normen aus. Die Kriterien der Fahreignung – im Sinne von Mindestanforderungen – müssen für alle gelten, unabhängig vom Alter.

Die Größe der Eichstichprobe von in Deutschland zugelassenen Testverfahren sollte nach Ansicht der Teilnehmer insgesamt mindestens 1.200 Personen umfassen, mindestens 200 Personen pro Altersgruppe. Die Gruppenaufteilung wurde nicht näher spezifiziert. Es wurde lediglich angeregt, die 30- bis 50-Jährigen in einer Gruppe zusammenzufassen und für höhere Altersgruppen eine feinere Gruppendifferenzierung vorzunehmen (z. B. 70-75 Jahre).

5) Validität und Validierung

Die Teilnehmer bewerten die Validierung der vorliegenden Testsysteme insgesamt als unzureichend. Wie Prof. Schmidt-Atzert betont, brauche man zunächst eine wissenschaftlich abgesicherte Konstruktvalidität, anschließend Kriteriumsvalidität. Es muss empirisch belegt sein, so Dr. Goldhammer, dass der Test geeignet ist, Fahrtaugliche von Fahruntauglichen zu trennen. Der Testautor muss hierzu Kennwerte (z. B. Zahl der Falsch-Positiven) angeben. Die Teilnehmer sind sich einig, dass man grundsätzlich von den Testautoren verlangen muss, bei den entwickelten psychometrischen Testverfahren zur Überprüfung der Fahreignung nachvollziehbare Kenngrößen zur Kriteriumsvalidität anzugeben.

Hier gebe es jedoch keinen „Königsweg“, den man von vorneherein spezifizieren kann. Die Kriteriumsvalidität sollte, ähnlich wie bei einem Peer-Review-Verfahren, von unabhängigen Gutachtern wissenschaftlich geprüft werden. Wenn die Gutachter unabhängig voneinander zum Schluss kommen, dass der Weg zur Ermittlung der Kriteriumsvalidität angemessen ist und die Kenngrößen aus wissenschaftlicher Sicht ausreichend hoch sind, soll die Kriteriumsvalidität als ausreichend betrachtet und der Test zugelassen werden.

Extremgruppenvergleiche sind im Rahmen einer Validierungsstudie nur bedingt geeignet. Wie SOMMER, ARENDASY, SCHUHFRIED & LITZENBERGER, (2005) gezeigt haben, können 89 % der Personen einer Stichprobe auf der Basis eines bestimmten Testbündels richtig als Unfälle (Personen mit mehr als einem Unfall) bzw. Nicht-Unfälle klassifiziert werden. Die Klassifizierungsrate bei einer Stichprobe, bestehend aus Nicht-Unfällen und Personen, die genau einen 1 Unfall haben, wurde bislang nicht angegeben.

Ein Großteil der Teilnehmer spricht sich für Fahrproben zur Validierung von Fahreignungstests aus. Einige Vorbehalte sind dabei anzumerken: (1) Aus dem Bereich der Demenzforschung kommt der Hinweis, Fahrproben müssen einen höheren Schwierigkeitsgrad besitzen, damit deren Ergebnisse höher mit neuropsychologischen Tests korrelieren. (2) SOMMER et al. (2006) haben herausgefunden, dass über 80 % der Personen, die eine standardisierte Fahrprobe absolviert haben, aufgrund einer bestimmten Anzahl von Tests richtig klassifiziert wurden zur Gruppe, die ein positives bzw. ein negatives Globalurteil über das Fahrverhalten bei einer standardisierten Fahrprobe erhalten haben. Problem: Die Urteilswerte reichten von 1-5. Erst wer einen Wert größer 3,33 hatte, bekam ein positives Globalurteil. Die Asymmetrie führt zur Extremgruppenbildung und damit zu eindeutigeren Ergebnissen.

Prof. Rudinger schlägt eine Kriteriumsvalidierung von Leistungsprofilen vor. Dabei würden nicht einzelne Leistungsdimensionen mit einem Kriterium in Beziehung gesetzt, sondern Profile mit einer Anzahl unterschiedlicher Leistungsdimensionen. Hierdurch, so folgert Herr Holte, würde man auch Ergebnisse darüber erhalten, welche Leistungsdefizite in einzelnen Bereichen kompensiert werden können. Zur Überprüfung der Konstruktvalidierung schlägt Prof. Rudinger zum Beispiel die Anwen-

derung der Generalisierbarkeitstheorie vor, ein messtheoretischer Ansatz, der im Gegensatz zur klassischen Testtheorie verschiedene Fehlerquellen gleichzeitig berücksichtigt. Der Multitrait-Multimethod-Ansatz bietet in der Diagnostik ebenfalls eine Möglichkeit, im Rahmen einer Konstruktvalidierung den Einfluss unterschiedlicher Einflussquellen (Traits, Methoden, Personen) zu berücksichtigen. Wie Prof. Schmidt-Atzert anmerkt, wird ein solcher Ansatz in der Fahreignungsdiagnostik nicht einfach zu realisieren sein und sollte daher nicht als Standardverfahren definiert werden.

Im Rahmen einer Kriteriumsvalidierung ist die Festlegung des Kriteriums von entscheidender Bedeutung und betrifft die Frage, welche Verhaltensweisen im Kontext realer Fahrsituationen Aussagen über die Fahreignung einer Person erlauben. Zur Klärung dieser Frage halten insbesondere Prof. Schmidt-Atzert, Prof. Rudinger sowie Dr. Poschadel es für erfolgversprechend, mit einer Anforderungsanalyse des Fahrverhaltens zu beginnen. Dabei sollte festgestellt werden müssen, welche Anforderungen für ein sicheres Autofahren erfüllt werden und welche kognitiven Fähigkeiten dafür mindestens vorliegen müssen. Diese Erkenntnisse sollten auch in die Entwicklung und Weiterentwicklung von Leistungstests einfließen.

6) Reliabilität

Die Reliabilität eines Tests zur Fahreignungsdiagnostik wird häufig ausschließlich durch das Cronbachs Alpha (interne Konsistenz) angegeben. Wie die Teilnehmer bekräftigen, ist dies nur ein Aspekt der Reliabilität. Die Teilnehmer fordern übereinstimmend für jeden zugelassenen Test die Angabe eines Stabilitätsmaßes (Retest-Reliabilität), um beurteilen zu können, wie sich die Zuverlässigkeit eines Tests über die Zeit verändert. Nur eine hohe Stabilität erlaubt auch sichere Prognosen. Eine solche Prognose, so Prof. Rudinger, sollte mindestens 8 Jahre halten. Das ist sehr unwahrscheinlich, wenn bereits der erste Stabilitätskoeffizient (1. Intervall) sehr niedrig ist. Wenn dieser z. B. bei .70 liegt, so Prof. Rudinger, dann wäre im 2. Intervall bereits ein Wert von nur .49 zu erwarten und im darauffolgenden Intervall ein Wert von .24 usw. Ein Stabilitätskoeffizient sollte mindestens für ein halbes Jahr angegeben werden, besser wäre für ein ganzes Jahr.

Wenn Tests wiederholt werden, muss grundsätzlich die Trainierbarkeit einer Aufgabe beachtet werden, darauf weist Herr Brenner-Hartmann hin. Es sei

daher erforderlich zu klären, wann und ob man bestimmte Tests in einer Testsituation wiederholen darf. Dazu ergänzt Prof. Schmidt-Atzert, dass der Testanwender wissen müsse, wie groß die Trainingseffekte sind. Diese Angaben gehörten ins Testmanual und müssten bei der Testkonstruktion untersucht werden.

7) Objektivität

Nach Ansicht aller Teilnehmer ist die Objektivität der derzeitigen psychometrischen Testverfahren weitgehend gegeben. In dieser Diskussion wurde der Einfluss technischer Gegebenheiten auf die Objektivität von Tests als relevant angesehen, jedoch nicht näher vertieft. Die Diskussion speziell über die Objektivität von Reaktionszeitmessungen wurde nicht geführt, da die BAST zu dieser Thematik zunächst die Ergebnisse dieses Projekts (siehe Ergebnisse in Kapitel 10) abwarten möchte.

Zur Gewährleistung der Objektivität hält Herr Brenner-Hartmann eine Schulung der Testanwender für wichtig. Testanwender müssten durch eine entsprechende Berufsausbildung dazu qualifiziert sein, wissenschaftliche Tests durchzuführen und auszuwerten, um die Testergebnisse kritisch würdigen zu können. Die Testanwender müssen wissen, welche der zahlreichen Kennwerte bedeutsam sind und wie sie interpretiert werden müssen. Derzeit besteht das Problem, dass die Objektivität eines Testverfahrens durch Anwender verfälscht werden kann, indem z. B. die Gewichtung von Subtests willkürlich gehandhabt wird oder Tests so lange wiederholt werden, bis ein gutes bzw. ausreichendes Prüfergebnis vorliegt.

Grundsätzlicher Konsens besteht in der Forderung, dass Testverfahren nur von dafür ausgebildeten Personen durchgeführt werden sollen, um eine fundierte Einschätzung und Interpretation der Testergebnisse zu gewährleisten. Die Testmanuale müssen übersichtlicher und verständlicher gestaltet sein, z. B. dass ein einzelnes Testergebnis richtig in den Zusammenhang der Normstichprobe eingeordnet werden kann, indem Normwerttabellen mitgeliefert werden und nicht nur vollautomatisierte „Endberichte“ bei computergestützten Messverfahren ausgedruckt werden.

8) Qualitätssicherung der Testentwicklung

Die Qualitätssicherung in der Begutachtung der Fahreignung wird durch die Akkreditierungsstelle

der BAST wahrgenommen. Eine Qualitätssicherung bei der Entwicklung von Tests zur Überprüfung der Fahreignung existiert jedoch gegenwärtig nicht. Die bestehende Praxis ist im Augenblick so, dass eine gutachterliche Stellungnahme von einem Experten aus der Verkehrssicherheitsszene eingeholt wird (beauftragt vom Testautor, Träger, Länder). Es sind für diese Begutachtung jedoch keine objektiven Kriterien spezifiziert worden, die nachweislich eingehalten werden müssen.

Die Teilnehmer fordern daher, dass die Prüfung neu beantragter psychometrischer Testverfahren nur von einer unabhängigen Instanz erfolgen soll. Diese Vorgehensweise gewährleiste die Einhaltung wissenschaftlicher Mindeststandards bei der Entwicklung der eingesetzten Testverfahren. Eine solche Instanz gibt es für Testverfahren zur Diagnose der Fahreignung bislang nicht. Es besteht ferner Konsens darüber, dass eine solche Prüfung nach vorher festgelegten Kriterien erfolgen und von mindestens zwei unabhängigen Gutachtern vorgenommen werden müsse. Hierzu wären u. a. folgende Fragen zu klären: Welche Vorgehensweise kann gewählt werden (Kompetenzen des Bundes/Kompetenzen der Länder)? Welche Bewertungskriterien sollen diesem Zulassungsverfahren zugrunde gelegt werden? Welche und wie viele Testbewerter/Gutachter sind erforderlich? Wie wird deren unabhängige Begutachtung gewährleistet? Welche Anforderungen sollen an die Testautoren gestellt werden? Prof. Schmidt-Atzert wies auf die Arbeit des Testkuratoriums hin, das ein solches Vorgehen bereits im Rahmen der Testbewertung umgesetzt hat. Es würden in der Regel zwei Gutachter genügen, die zuerst unabhängig voneinander einen Test begutachten und dann ihre Gutachten zusammenführen. Dieses Vorgehen wird derzeit im Rahmen der DIN-Norm 33430 bei der Überprüfung von Tests zu berufsbezogenen Eignungsprüfungen praktiziert.

8.6 Optimierungsbedarf

Die in der Fahreignungsdiagnostik derzeit eingesetzten psychometrischen Testsysteme weisen nach Ansicht von Experten eine Reihe methodischer Schwächen auf, die vor allem die Gütekriterien betreffen. Insgesamt werden die von der FeV geforderten wissenschaftlichen Anforderungen an einen Fahreignungstest nicht voll umfänglich erfüllt, worin ein Optimierungsbedarf gesehen wird. Beim Expertentreffen wurde konsensual deutlich, dass

eine Optimierungsstrategie in zwei Schritten erfolgen sollte:

8.6.1 Kurzfristige Strategie

In einem ersten Schritt sollte darauf hingewirkt werden, dass die basalen wissenschaftlichen Testgütekriterien – wie von der FeV gefordert – von den in Deutschland eingesetzten Testverfahren im vollen Umfang erfüllt werden.

- Das betrifft die Eichung der psychometrischen Testverfahren. Diese sollten nachträglich an ausreichend großen (min. 1.200 Personen) repräsentativen Stichproben geeicht werden.
- Die Objektivität der Testanwendung muss gewährleistet sein. Es muss ausgeschlossen werden, dass die Objektivität eines Testverfahrens durch Anwender verfälscht wird, indem z. B. die Gewichtung von Subtests willkürlich gehandhabt wird oder Tests wiederholt werden, bis ein gutes/ausreichendes Prüfergebnis vorliegt.
- Alle Reliabilitätsmaße müssen nachvollziehbar dokumentiert werden, auch an welcher Stichprobe (welchen Stichproben) sie in welchem zeitlichen Rahmen erhoben wurden.
- Es müssen Stabilitätsmaße angegeben werden um nachvollziehen zu können, welche prognostische Güte die einzelnen Testverfahren haben (einschließlich der Subtests).
- Die Konstruktvalidität muss angegeben werden.
- Gleiches gilt für die Kriteriumsvalidität. Es wird angeregt, die Kriteriumsvalidität anhand einer „angereicherten Fahrprobe“ zu überprüfen. Eine Idee ist es, Leistungsprofile für eine Validierung heranzuziehen, also zu prüfen, inwieweit bestimmte Leistungsprofile geeignete Prädiktoren des Kriteriums sind.
- Zur Bewertung von Testsystemen müssen Kriterien aufgestellt werden, die von den zugelassenen Testverfahren als Mindeststandards einzuhalten sind. Dabei kann der in dem Projekt entwickelte Kriterienkatalog zunächst als Arbeitsgrundlage dienen.
- Einrichtung einer unabhängigen wissenschaftlichen Instanz bestehend aus unabhängigen wissenschaftlichen Gutachtern (vielleicht in einem „Blind peer-review-Verfahren“), die die Einhaltung dieser Kriterien überprüfen. Eine erste Vorlage für einen Kriterienkatalog könnte der im

Projekt entwickelte Kriterienkatalog sein, der entsprechend den Ergebnissen der Experten-Gruppe abgeändert werden kann.

Erst wenn diese Voraussetzungen erfüllt sind, sollte ein psychometrisches Testverfahren zur Überprüfung der Fahreignung (weiterhin) zugelassen werden dürfen.

8.6.2 Langfristige Strategie

Längerfristig müssen aus wissenschaftlicher Perspektive die in der FeV, Anlage 5, genannten wissenschaftlichen Konstrukte (Belastbarkeit, Orientierungsleistung, Konzentrationsleistung, Aufmerksamkeitsleistung, Reaktionsfähigkeit) grundsätzlich überarbeitet werden, da sie nicht dem Stand der aktuellen Forschung entsprechen und somit auch nur unscharf operationalisierbar sind. Darauf aufbauend sollen Tests zur Überprüfung der neueren wissenschaftlichen Konstrukte entwickelt oder eingesetzt werden, sofern sie schon existieren. Aus dieser Zielsetzung ergibt sich ein erheblicher Forschungsbedarf. Auch langfristig sollte ein unabhängiges wissenschaftliches Gremium die Güte der zugelassenen Testverfahren regelmäßig überprüfen.

Die theoretischen Grundlagen zur Konstruktion von psychometrischen Tests zur Fahreignungsdiagnostik sind derzeit defizitär. Eine theoretische Fundierung der Testentwicklung für psychometrische Testverfahren zur Überprüfung der Fahreignung wird übereinstimmend als dringend erforderlich gesehen. Es muss ein „Neuer Kanon“ mit aktuellen wissenschaftlichen Konstrukten erarbeitet werden, der anschließend in die FeV, Anlage 5, aufgenommen werden kann. Dieser Kanon sollte zunächst „Aufmerksamkeit plus“ genannt werden. Das heißt, es sollte zunächst vom kognitiven Modell der Aufmerksamkeit ausgegangen werden, in das weitere fahreignungsrelevante Konstrukte integriert werden können.

Hierfür scheint es erfolgsversprechend zu sein, kriteriumsorientiert vorzugehen: Beginnend mit einer Anforderungsanalyse des Fahrverhaltens im Realverkehr ist festzustellen, welche kognitiven Fähigkeiten mindestens vorliegen müssen, um die Fahraufgabe zu erfüllen, und anschließend die passenden Tests zu suchen bzw. zu entwickeln.

Langfristig sollten die in der Fahreignungsdiagnostik angewendeten Entscheidungskriterien (Prozentränge) durch „echte Kriterien“ ersetzt werden, also durch absolute Kennwerte, die auf dem Wis-

sen über die Zusammenhänge zwischen Leistung und Fahreignung basieren. Diese Zusammenhänge sind noch nicht ausreichend erforscht.

9 Die Anforderungsbereiche der FeV unter neueren wissenschaftlichen aufmerksamkeits-theoretischen Aspekten

9.1 Aktuelle Aufmerksamkeitstheorien

Wie in Kapitel 2 dargestellt, stellt „Aufmerksamkeit“ ein heterogenes Konstrukt dar, das sowohl unter behavioraler als auch funktionell-bildgebender Betrachtungsweise unterschiedliche Aspekte aufweist, die weit über die in der FeV genannten Anforderungsbereiche hinaus gehen. Zusammenfassend lassen sich folgende Aspekte unterscheiden:

Intensitätsaspekt

- Alertness (allgemeines Aktivierungsniveau),
- Daueraufmerksamkeit/Vigilanz.

Selektive, räumliche Aufmerksamkeit

- Verdeckte Aufmerksamkeitsverschiebung (häufig im Vorfeld von Blickbewegungen),
- offene Aufmerksamkeitsverschiebung (mit Blickbewegungen),
- visuelles Suchen,
- Überblicksgewinnung in Verkehrssituationen.

Selektive nicht-räumliche Aufmerksamkeit (exekutive Aufmerksamkeit)

- Fokussierung auf bestimmte Zeitpunkte/Objekte/Reiz Aspekte,
- Inhibition von Reaktionen (zur Vermeidung von erhöhter Ablenkbarkeit),
- crossmodale Integration,
- Aufmerksamkeitsteilung,
- Strategie/Flexibilität.

Bei der selektiven Aufmerksamkeit ist zudem eine modalitätsspezifische Verarbeitung hervorzuheben. Je nach Fokus der Aufmerksamkeit wird die Verar-

beitung von sensorischer Information in den entsprechenden primär-sensorischen Hirnarealen erleichtert (SOMERS, DALE, SEIFFERT & TOOTELL, 1999). Für die Fahreignung wird jedoch die visuelle Verarbeitung eindeutig im Vordergrund stehen.

Die 5 Anforderungsbereiche der FeV lassen sich nur in Ansätzen in diese aktuelle Konzeption der Aufmerksamkeit integrieren (vgl. Tabelle 1).

9.2 Welche Aufmerksamkeitsaspekte sind aus empirischer Sicht wichtig für die Fahreignung?

In Kapitel 3 werden Ergebnisse der Literaturanalyse zum Zusammenhang zwischen Aufmerksamkeitstests und dem Kriterium der Fahreignung referiert. Dabei ergibt sich kein eindeutiges Resultat zu Gunsten eines oder mehrerer Tests. Allerdings finden sich insgesamt etwas höhere Korrelationen zwischen Tests, die visuell-räumliche Aufmerksamkeitskomponenten erfassen, und der Fahreignung. Hierbei sind vor allem die Tests Wiener Determinationstest, Visuelles Scanning (TAP-M), UFOV und Trailmaking (insbesondere Form B) hervorzuheben. Dies könnte somit darauf hindeuten, dass der visuell-räumlichen Aufmerksamkeit eine wichtige Rolle bei der Fahreignung zukommt. Gleichwohl ist festzustellen, dass die prädiktive Validität der psychometrischen Tests in den genannten Studien sich als allenfalls mittelmäßig erweist, was sowohl auf mangelnde Reliabilität oder Validität der Testverfahren, jedoch auch durchaus auf die Art der Messung des Kriteriums zurückgeführt werden kann. Weiterhin wurden die Korrelationskoeffizienten der Subtests des ART 2020 mit der Fahrprobe leider nicht angegeben, daher ist hier eine Beurteilung der Vorhersagekraft der Untertests nicht möglich. Bei Patientenpopulationen (Demenz, Parkinson) sind häufig auch Gedächtnistests ausschlaggebend (Rey-Osterreith, MMSE). Hier kann jedoch davon ausgegangen werden, dass eher die Schwere der Krankheit und die hiermit einhergehende Beeinträchtigung erfasst wird; es ist weniger plausibel anzunehmen, die Merkfähigkeit würde einen signifikanten Beitrag zur Fahreignung erbringen.

9.3 Zuordnung der kriteriumsvalidierten Testverfahren zu den Anforderungsbereichen der FeV

Eine empirische Zuordnung erfolgt aufgrund der fehlenden Einzeldaten nicht für den ART 2020, hier sei allerdings auf die Publikation von BUKASA (2003) verwiesen, in der die Autoren eine Zuordnung vornehmen. Zusätzlich gehen theoretische Überlegungen über den Testaufbau in die Einschätzung ein. Die anderen (Unter-)Testverfahren, die hier zugeordnet werden, haben sich in den oben genannten Studien als aussagekräftig gezeigt.

Die verschiedenen Testverfahren können nach Meinung der Autoren folgendermaßen in die Anforderungsbereiche der FeV eingeordnet werden (siehe Tabelle 4).

Somit könnten alle Anforderungsbereiche in ihrer derzeitigen Festlegung mit den vorhandenen Testverfahren überprüft werden. Dennoch muss festgehalten werden, dass die Anforderungsbereiche einer Überarbeitung bedürfen und dass die Überprüfung der Testverfahren, wie oben erläutert, an deutlich verbesserten Standards erfolgen muss (siehe insbesondere Kapitel 8.5 und Kapitel 8.6). Die in Tabelle 4 erbrachte Zuordnung ist somit als der Versuch einer derzeitigen möglichen Interimslösung zu verstehen, bis die erforderlichen Verbesserungen auf Seiten der Anforderungsbereiche, der Testverfahren und deren Validierung durchgeführt worden sind.

Anforderungsbereich	Computergestützte/apparative Tests
Belastbarkeit	Wiener Determinationsgerät ART 2020 – RST3
Orientierung	TAP-Visuelles Scannen TAVT 2 Wiener Determinationsgerät ART 2020 – LL5, TT15
Konzentration	Wiener Reaktionstest (Ton-Licht) ART 2020 – Q1
Aufmerksamkeit	TAVT 2 TAPK-Geteilte Aufmerksamkeit TAPK-Reaktionswechsel ART 2020 – PVT, Q1
Reaktionsfähigkeit	Wiener Reaktionsgerät (Ton bzw. Licht) Wiener Determinationstest ART 2020 – DR2, FAT, SET3

Tab. 4: Mögliche Zuordnung der Testverfahren zu den Anforderungsbereichen

9.4 Vorschlag zur Schärfung der derzeit bestehenden Anforderungsbereiche

Ausgehend von theoretischen und empirischen Befunden werden folgende Anforderungsbereiche vorgeschlagen:

- Allgemeine Reaktionsbereitschaft (Alertness; entspricht weitgehend dem aktuellen Bereich „Reaktionsfähigkeit“).
- Räumliche Aufmerksamkeitsausrichtung (ähnlich dem jetzigen Bereich „Orientierung“).
- Aufmerksamkeitsteilung (entspricht teilweise dem jetzigen Bereich „Aufmerksamkeit“).
- Fokussierte Aufmerksamkeit (entspricht teilweise dem Bereich „Konzentration“ und beinhaltet die Fähigkeit, sich nicht durch störende visuelle und/oder auditive Reize ablenken zu lassen).
- Daueraufmerksamkeit (im Sinne von „Belastbarkeit“).

Dabei wäre es wichtig, präzisere Operationalisierungen dieser neu formulierten Anforderungsbereiche anzugeben und vor allem auf spezifische Situationen im Straßenverkehr zu beziehen.

9.5 Vorschlag einer Testbatterie zur Fahreignungsprüfung

Die derzeit verwendeten Testverfahren werden weitestgehend, auch nach einer Schärfung der Anforderungsbereiche, diesen „neuen“ Bereichen zugeordnet werden können. Dies betrifft Konstrukte, die nur einer geringen Korrektur bedürfen (Reaktionsfähigkeit und Orientierung/räumliche Aufmerksamkeit). Hier können z. B. die entsprechenden Subtests des ART 2020, des Wiener Determinationstests oder entsprechende Subtests der TAP-M angewendet werden (s. Tabelle 4). Bezüglich der drei anderen Bereiche muss jedoch nach einer genauen Ausformulierung die Passung der Testverfahren erneut überprüft werden. Gerade die Frage der Ablenkbarkeit bedarf genauerer Betrachtung, während Daueraufmerksamkeit möglicherweise durch den Wiener Determinationstest oder das Wiener Testsystem-Daueraufmerksamkeit oder TAM-P Alertness erfasst und die Aufmerksamkeitsteilung vermutlich durch TAVT2 oder TAP-M Geteilte Aufmerksamkeit abgedeckt wird.

10 Überprüfung der Reaktionszeitmessung eines computergestützten psychometrischen Testverfahrens

10.1 Einleitung

Die Fahrerlaubnis-Verordnung (FeV), Anlage 5, legt Rahmenbedingungen für die Begutachtung von Bewerbern um die Erteilung oder Verlängerung der Führerscheinklassen D, D1, DE, DE1 sowie einer Fahrerlaubnis zur Fahrgastbeförderung fest. Derzeit fehlen konkrete Bestimmungen, welche psychometrische Verfahren – und in welchem Umfang – Bestandteil der Begutachtung nach § 11 Abs. 9 und § 48 Abs. 4 und 5 FeV sein sollen. Festgelegt werden allein die Bereiche, die begutachtet werden sollen. Nicht festgelegt wird, mit welchen konkreten Verfahren diese psychologischen Aspekte der Fahreignung untersucht werden sollen.

Derzeit legen die Bundesländer fest, welche (computergestützten) Testverfahren in den jeweiligen Bundesländern zur Überprüfung der genannten Bereiche zugelassen sind. In einem Teilprojekt, das Gegenstand des letzten Kapitels des Gesamtforschungsberichtes ist, wird ein in allen Bundesländern Deutschlands zugelassenes computergestütztes Testverfahren zur Überprüfung der psychometrischen Leistungsprüfung dahingehend untersucht, ob die vom Testprogramm ermittelten Reaktionszeiten für den Bereich „Reaktionsfähigkeit“ auch mit objektiv ermittelten Reaktionszeiten übereinstimmen und somit die (technische) Objektivität zur Ermittlung von Reaktionszeiten uneingeschränkt gewährleistet ist.

10.2 Problemstellung

Die Anlage 5 der Fahrerlaubnisverordnung (FeV) nennt 5 Anforderungen an die psychische Leistungsfähigkeit: Es sind dies (a) Belastbarkeit, (b) Orientierungsleistung, (c) Konzentrationsleistung, (d) Aufmerksamkeitsleistung und (e) Reaktionsfähigkeit. Diese Anforderungsbereiche bezeichnen unterschiedliche Aufmerksamkeitsaspekte, die zuverlässig und valide von den jeweiligen Testverfahren im Rahmen der Fahreignungsprüfung erfasst werden müssen.

In den Begutachtungs-Leitlinien zur Krafftfahreignung (Bundesanstalt für Straßenwesen, 2000) wer-

den darüber hinaus Grenzwerte festgelegt, die bei der Ermittlung der Leistungswerte in den verschiedenen Aufmerksamkeitsbereichen mindestens erreicht werden müssen: „Zweifel an der psychischen Leistungsfähigkeit können sich ergeben wegen einer Minderung der optischen Orientierung, der Konzentrationsfähigkeit, der Aufmerksamkeit, der Reaktionsfähigkeit und der Belastbarkeit. [...] Die psychologische Leistungsfähigkeit wird mit geeigneten, objektivierbaren psychologischen Testverfahren untersucht. Ausschlaggebend ist, ob die Mindestanforderungen erfüllt werden“ (BASt, 2000, S. 16). Die Mindestanforderungen werden im Wesentlichen wie folgt definiert: „Die Zweifel können in der Regel als ausgeräumt gelten, wenn sich eine der folgenden Feststellungen treffen lässt:

Gruppe 1 [das sind im wesentlichen Auto- und Motorradfahrer, Anm. d. Verf.]

Der Prozentrang 16 wurde, bezogen auf altersunabhängige Normwerte, in allen Leistungstests erreicht oder überschritten“ (ebd., S. 17).

Im Weiteren wird eingeräumt, dass es Ausnahmen von dieser Regel geben kann,

- wenn Einzelleistungen in einem Bereich in anderen Bereichen der psychischen Leistungsfähigkeit kompensiert werden können oder
- wenn schlechte Leistungen als situationsbedingt anzusehen sind oder
- wenn in anderen Ergänzungsverfahren eine bessere Leistung nachgewiesen werden kann oder
- sich der Fahrer in der Praxis bewährt hat (Fahrverhaltensprobe) und keine Hinweise auf verkehrsmedizinisch relevante Mängel vorliegen (vgl. ebd.).

Für die Gruppe 2 der Fahrer, das sind im Wesentlichen Lkw-Fahrer, Busfahrer und Fahrer, die eine Erlaubnis zur Fahrgastbeförderung besitzen, gelten strengere Maßstäbe in Bezug auf die psychische Leistungsfähigkeit: „Es gelten sinngemäß die Ausführungen zur Gruppe 1. Darüber hinaus gilt die erhöhte Anforderung, dass in der Mehrzahl der eingesetzten Verfahren der Prozentrang 33 – gemessen an altersunabhängigen Normwerten – erreicht oder überschritten werden muss, dass aber der Prozentrang 16 in den relevanten Verfahren ausnahmslos erreicht werden muss.

Hiervon kann nur abgesehen werden, wenn in einzelnen Untertests bei Abweichungen nach unten

Kompensationsmöglichkeiten gegeben sind. Andererseits muss sichergestellt werden, dass eine Kumulation ausgeschlossen ist (siehe Gruppe 1).

In Zweifelsfällen ist eine Fahrverhaltensprobe durch den psychologischen Gutachter vorzunehmen (Gruppe 1 und 2)“ (ebd. S. 17).

In den Begutachtungs-Leitlinien zur Kraftfahrereignung wird also recht deutlich, dass ein Unterschreiten der geforderten Leistungen in Teilbereichen der psychologischen Leistungsfähigkeit erhebliche Konsequenzen für den Führerscheinwerb respektive den Fortbestand der Fahrerlaubnis haben kann. Insofern muss die Objektivität der eingesetzten Testverfahren uneingeschränkt gewährleistet sein.

Die Zuverlässigkeit und Validität psychologischer Testprogramme, die in der Fahreignungsdiagnostik zur Anwendung kommen, sind also abhängig von einer objektiven und reliablen Erfassung der für das Autofahren relevanten Leistungsparameter. Da in der Fahreignungsdiagnostik computerbasierte Testverfahren zum Einsatz kommen, für die eine entsprechende Zulassung seitens der Bundesländer existiert, besteht die berechtigte Frage, inwieweit unterschiedliche Hard- und Softwarekonstellationen, wie sie in der Praxis in der Regel üblich sind, die Objektivität der Reaktionszeitmessungen beeinflussen können.

Im Rahmen einer experimentellen Studie wurde erforscht, ob unterschiedliche Hard- und Softwarekonstellationen tatsächlich Einfluss auf das vom Computerprogramm ermittelte Ergebnis der Reaktionszeitmessung für das Konstrukt „Reaktionsfähigkeit“ nach FeV haben können. Sollten sich Differenzen zur externen, objektiven Messung ergeben, wären die Objektivität (eventuell auch die Reliabilität) und damit die Validität der computergestützten Reaktionszeitmessung für dieses Konstrukt unmittelbar berührt.

Für die experimentelle Studie wurde eines der in allen Bundesländern zugelassenen computergestützten Testverfahren untersucht.

10.3 Ausgangslage

Die in Anlage 5 der Fahrerlaubnisverordnung (FeV) genannten Anforderungen an die psychische Leistungsfähigkeit beinhalten auch eine Untersuchung der Reaktionsfähigkeit, bei der schon Unterschiede im Millisekundenbereich vermutlich Auswirkungen

auf den entsprechenden Prozentrangwert haben können. Deshalb muss die Reaktionszeit millisekundengenau erfasst werden, um die vorgeschriebene Unterteilung in verschiedene Prozentränge und einen Vergleich mit Normtabellen leisten zu können. In eigenen Untersuchungen im Bereich der Grundlagenforschung zur Erfassung Ereigniskorrelierter Potenziale (EKP) beschäftigt sich die Projektgruppe „Altern und ZNS-Veränderungen“ (ehemals Projektgruppe „Kognitive Neurophysiologie“) am Institut für Arbeitsphysiologie an der Universität Dortmund (IfADo) seit etlichen Jahren mit der Problematik der computergestützten zeitgenauen Präsentation von auditiven und visuellen Reizen und der damit verbundenen millisekundengenauen Erfassung der Reaktionszeiten von Probanden. Bei Ablenkungsaufgaben oder Wahlreaktionenaufgaben beispielsweise liegen die zu erwartenden Verhaltensunterschiede im Bereich von 10 Millisekunden (ms) bis zu 100 ms. Parallel zu diesen Verhaltenaufgaben wird das EEG (Elektroenzephalogramm) abgeleitet und mit Hilfe der EKP werden so Verhaltensunterschiede objektiv dargestellt und lokalisiert. Zufällige oder systematische Fehlmessungen im Millisekundenbereich, die durch eine ungünstige Soft- und Hardwarekonstellation bedingt wären, könnten tatsächliche Effekte entweder maskieren oder Messergebnisse liefern, die – objektiv betrachtet – einen Artefakt darstellen. Um EKP aus dem EEG zu generieren, ist es zwingend erforderlich, absolut zeitsynchron (das heißt geeicht) und millisekundengenau alle Ereignisse² der Laboraufgaben zu markieren und elektronisch zu erfassen, sodass beim Averaging-Verfahren die EKP nicht jittern (also nicht zeitlich fluktuieren in Relation zur zeitlichen Präsentation des Target-Reizes auf einem Bildschirm) und somit „verschmieren“. Derart „verschmierte“ Daten wären nicht bzw. nur sehr ungenau auswertbar und daher in der Grundlagenforschung nicht verwertbar. Das absolut zeitgenaue Präsentieren und Markieren von auditiven sowie visuellen Reizen sind für die STA-Analyse (Stimulus Triggered Average) der EKP wichtig, das genaue Markieren der RT ist zwingende Voraussetzung für

die RTA-Analyse (Reaction Triggered Average). Um diese verschiedenen Anforderungen sicherzustellen, werden in der Arbeitsgruppe Betriebssysteme verwandt, die keine so genannten „Multitasksysteme“ sind. Die Steuerprogramme sind selbst programmiert und greifen außerdem auf selbst entwickelte PC-Hardware zurück, um von der Taktrate eines PCs unabhängig zu bleiben. Eine Überprüfung der Funktionstüchtigkeit der Steuerprogramme und deren Eichung werden mit externer Hardware durchgeführt. Diese Eichung wird bei jeder Programmierung einer Laboraufgabe und bei jedem Versuchsaufbau standardmäßig durchgeführt, um die Qualität der wissenschaftlichen Messungen garantieren zu können. Wenn nötig, werden synthetische Reize als Reizmuster programmiert und in die EEG-Aufzeichnungen mit eingespeist, um bei der Auswertung der Daten fortlaufend im Bereich von 0,1 ms messen und auswerten zu können.

Die in den Bundesländern zugelassenen Tests zur computergestützten Reaktionszeitmessung bei der psychometrischen Fahreignungsdiagnostik wurden von den Testautoren für so genannte Multiuser-Multitask-Betriebssysteme entwickelt (Windows 98, Windows 2000 sowie Windows XP). In Bezug auf die Problematik einer millisekundengenauen Messung zeichnen sich Multitasksysteme vor allem dadurch aus, dass parallel laufende Prozesse im Hauptprozessor des Computers je nach Priorität durch Hintergrundprozesse jederzeit – jedoch nicht vorhersehbar – unterbrochen werden können, ohne dass der eigentliche Computernutzer dieses Eingreifen steuern könnte. Damit wird eine millisekundengenaue Messung von Reaktionszeiten durch die Architektur des Betriebssystems von Anfang an erheblich erschwert. Die Dauer solcher Unterbrechungen ist von der Rechengeschwindigkeit des Computersystems (Taktfrequenz) abhängig. Die Anzahl sowie die Prioritäten von unterbrechenden Prozessen hängen von der Software-Konstellation eines Computers ab. Als Beispiel sei ein installierter und aktivierter Virens scanner genannt, der zeit- oder auch objektgesteuert das System „ausbremsen“ und damit auch Messergebnisse verfälschen kann, indem laufende Prozesse z. B. einer unvorhergesehenen Virenüberprüfung unterzogen werden können. Einen erheblichen (und in der Regel noch größeren) Einfluss auf die Möglichkeit der millisekundengenauen Messung von Reaktionszeiten üben die Art und Weise der Darstellung der Zielreize (Targets) auf einem Computerbildschirm aus. Zwischen dem im Computerprogramm zeitgenau

² Das sind im Wesentlichen die zeitgenaue Aufzeichnung des zeitlichen Erscheinens des Hinweisreizes auf dem Bildschirm, des Target-Reizes auf dem Bildschirm, die Reaktion des Probanden an einem Device (Eingabegerät) und des EKP, das wie eine EEG-Kurve aussieht, aber an verschiedenen Stellen des Kopfes des Probanden simultan abgeleitet wird (in der Regel 64-Kanal-Aufzeichnungen).

programmierten Target und dem auf dem Bildschirm tatsächlich erscheinenden Target können zeitliche Unterschiede im Millisekundenbereich liegen, die einen Einfluss auf die Genauigkeit der Messergebnisse haben können, wenn nicht millisekundengenau synchronisiert wurde: Die Informationen über das auf dem Bildschirm darzustellende Target werden vom Computerprogramm zunächst in den Bildschirmspeicher der Grafikkarte geschrieben. Dieser Bildschirmspeicher wird zyklisch seriell ausgelesen, um ein Abbild auf dem Monitor zu erzeugen (Darstellung des Targets). Dieses zyklische Auslesen geschieht mit einer bestimmten Frequenz, der Bildwiederholfrequenz. Die Bildwiederholfrequenz kann bei PC-Systemen unterschiedlich eingestellt werden, abhängig von der Güte der Grafikkarte und des verwendeten Monitors. Übliche Werte liegen im Bereich zwischen 75 Hz und 100 Hz (13,33 ms bis 10,00 ms). Wird also programmtechnisch nicht automatisch mit der Bildwiederholfrequenz der jeweiligen aktuellen Soft- und Hardwarekonfiguration synchronisiert, können sich Darbietungsfehler im Bereich der Bildwiederholfrequenz einschleichen (10 ms bis 13 ms) und zu entsprechend großen Messfehlern führen. Das Zusammenspiel von Graphikkarte, Bildwiederholfrequenz und verwendetem Monitor stellt eine mögliche der wesentlichen Ursachen für Messfehler dar. Neben den genannten Hauptursachen kann die Erfassung der Reaktion des Probanden einen Fehler verursachen. Dies ist abhängig von dem verwendeten Response-Device („Reaktionserfassungs-Vorrichtung“). Maus und Tastatur übertragen ihre Daten seriell (also mit endlicher Genauigkeit) an den PC, wobei im PC der Prozess für die Auswertung der übertragenen Daten aktiviert wird und dann die Reaktionszeit gestoppt werden kann. Vielfach wird deshalb ein spezielles externes Response-Device eingesetzt, um den möglichen Fehler so gering wie möglich zu halten. Das Response-Device wird in der Regel am Printerport angeschlossen (bei neueren Entwicklungen auch am USB-Eingang des Computers) und die Reaktionsinformation im Pollingverfahren³ abgefragt. Neben den beschriebenen Einzelproblemen können die beschriebenen möglichen Fehlerquellen auch in Wechselwirkung

miteinander treten, wodurch im Einzelfall erhebliche Messwertfehler im Bereich von bis zu 100 ms – ebenfalls abhängig von der jeweiligen Soft- und Hardwarekonstellation – auftreten können, die die Größenordnung der Einzelmessfehler vielfach überschreiten, jedoch nicht vorhersagbar sind.

Durch die in diesem Kapitel beschriebenen technisch möglichen Einflüsse auf eine millisekundengenaue Erfassung von Reaktionszeiten können sich Unterschiede zwischen den durch ein computergestütztes Messverfahren ermittelten Reaktionszeiten und den objektiven Reaktionszeiten in einer Größenordnung ergeben, die einen Einfluss auf die richtige Zuordnung zu den Prozentrangwerten im Anforderungsbereich „Reaktionsfähigkeit“ nach FeV haben.

10.4 Methodisches Vorgehen/Untersuchungsdesign

Mit dem Forschungsprojekt sollen Abweichungen zwischen den von einem computergestützten Test ermittelten Reaktionszeiten und objektiv ermittelten Reaktionszeiten erhoben werden. Insbesondere wird der Frage nachgegangen, ob unterschiedliche Systemkonfigurationen, auf denen der psychometrische Test installiert ist, einen Einfluss auf das vom Computertest ermittelte psychologische Leistungsmerkmal „Reaktionsfähigkeit“ haben.

Hierfür wurde eines der in Deutschland in allen Bundesländern zugelassenen computergestützten Testverfahren zur Ermittlung der psychologischen Leistungsfähigkeit einer Zeitprüfung unterzogen.

Die Messergebnisse werden in anonymisierter Form dargestellt, da in der Testreihe nur der Frage nachgegangen wird, ob Fehlmessungen, die die Objektivität signifikant beeinflussen können, für eines der in der Praxis eingesetzten Testverfahren möglich sind.

Um möglichst realistische Ergebnisse zu erhalten, die auch dem Einsatz der Testsysteme in der Praxis entsprechen, wurden die Versuchsreihen mit einem „echten“ Probanden durchgeführt. Hiermit wurde auch einer möglichen Wechselwirkung zwischen einem „elektronischen“ Probanden und dem Messsystem entgegengewirkt. Es sollte so verhindert werden, dass die Komponenten „computergestützter Test“ und „elektronischer Proband“ in irgendeiner Weise in Wechselwirkung treten und das System zu „schwingen“ beginnt.

³ Polling bezeichnet die Methode, den Status eines nicht beeinflussbaren Geräts aus Hard- oder Software mittels zyklischen Abfragens (wiederum abhängig von der internen Arbeitsfrequenz des Computers) zu ermitteln, die auch Sendeaufrufe genannt werden.

Um die Genauigkeit der Prozentwerteinteilung des untersuchten Messsystems abschätzen zu können, wurde vor Abschluss der Untersuchungen eine weitere Untersuchungsreihe mit einem elektronischen Probanden durchgeführt. Hierbei wurde ermittelt, wie sich der Messfehler des untersuchten computergestützten Messverfahrens (einer einzigen Konfiguration) auf die Zuteilung zu Prozentrangwerten in sensiblen Bereichen (16. Perzentil, 33. Perzentil, siehe Kapitel 10.2) auswirken kann. Dabei wurde die voreingestellte Reaktionszeit des elektronischen Probanden in den sensiblen Bereichen in 1 ms-Schritten variiert.

Als wesentliche Einflussgrößen für Fehlmessungen können (neben Programmierfehlern) folgende Komponenten eines Computersystems vermutet werden:

10.4.1 Hardware

- **Das PC-System**

Es wurden drei verschiedene PC-Systeme in die Versuchsreihe aufgenommen. Eine Konfiguration mit einem Standard PC mit derzeit gängigen Anwendungen. Eine Konfiguration mit einem neuen PC mit technischem Stand 10/2006. Eine Konfiguration mit einem Standard-Laptop.

- **Monitore**

Es wurden drei verschiedene Monitorkonfigurationen getestet:

Samsung CRT Monitor Samtron 210 Plus mit 75 Hz,

Samsung CRT Monitor Samtron 210 Plus mit 100 Hz,

ViewSonic LCD-Monitor VX924 mit 60 Hz.

10.4.2 Software

- **Betriebssysteme**

Es wurden drei verschiedene Betriebssysteme getestet, die zum Zeitpunkt des Beginns des Forschungsprojektes noch immer sehr verbreitet waren:

Windows 98 mit letztem aktuellstem Service-Pack (zur Lauffähigkeit dieses Betriebssystems mehr im Kapitel 10.5),

Windows 2000 mit aktuellstem Service-Pack und

Windows XP.

- **Anwendungen**

Es wurden zwei verschiedene Software-Systemkonfigurationen getestet:

Rechnersystem mit Minimal-Konfiguration „raw“ (das heißt mit von der CD neu installiertem Betriebssystem, zusätzlich mit dem oben beschriebenen und von Microsoft empfohlenen Service-Pack für Windows 2000) und

Rechnersystem mit Standard Office-Anwendungen (Word, Excel usw. incl. Virens Scanner).

10.4.3 Forschungsdesignplan

Damit ergibt sich grundsätzlich ein 3X3X3X2-Design: 3 PC-Systeme x 3 Monitore x 3 Betriebssysteme x 2 Softwarekonfigurationen. Somit ergaben sich für die Planung 54 Konfigurationen (zu nicht lauffähigen Konfigurationen mehr im nächsten Kapitel).

10.5 Versuchsaufbau

Die Hauptversuche, mit denen auch die Endergebnisse ermittelt wurden, wurden mit einer aktuellen Version des Testsystems durchgeführt. Die Vorversuche zum Test des Versuchsaufbaus wurden mit einer früheren – nicht aktuellen – Version durchgeführt.

Die Präsentation des Targets wurde mit Hilfe einer lichtempfindlichen Leuchtdiode direkt am Monitor abgegriffen und als Signal in ein externes Messwert-Erfassungssystem eingespeist, das in der Arbeitsgruppe seit vielen Jahren eingesetzt wird.

Das externe Erfassungssystem bestand hardwareseitig aus einem NeuroScan-64-Kanal-EEG-AD-Card System mit der Erfassungssoftware Acquire 4.0 (der Firma NeuroScan Inc, Texas), einem verbreiteten wissenschaftlichen medizinischen Standardgerät zur Erfassung von EEG-Daten.

Die Auswertung der Daten erfolgte mit der Software Analyzer 1.05 der Firma Brain Products GmbH sowie mit in der Arbeitsgruppe programmierten Hilfs- und Automatisierungsprogrammen.

Das Response-Device (das zum Lieferumfang des psychometrischen Testsystems gehört) wurde hardwareseitig direkt „angezapft“ (das heißt, dass der elektronische Impuls, der bei einer Reaktion durch den Druck auf einen Taster entsteht, direkt am Eingabegerät abgegriffen wurde, indem dort ein weiteres Kabel angelötet wurde) und das Signal ebenfalls in das Messwert-Erfassungssystem eingespeist. Zur erweiterten Kontrolle wurde als drittes Signal ein Rechtecksignal (1 kHz, das entspricht 1 ms) aus dem geeichten Funktionsgenerator mit erfasst.

Als Funktionsgenerator kam das Modell Agilent 33120A der Fa. Hewlett-Packard zum Einsatz, das von der Arbeitsgruppe in eigenen Untersuchungen regelmäßig verwendet wird. Die mögliche Abweichung bei der Frequenzgenerierung beträgt 10 ppm (parts per million) in 90 Tagen (AGILENT TECHNOLOGIES, INC. [Hrsg], 2002). Das Gerät ist ein verbreitetes medizinisches Standardgerät und genügt wissenschaftlichen Ansprüchen.

Die analog erhobenen Daten wurden kontinuierlich während des gesamten Versuches mit einer Taktrate von 10 kHz (0,1 msec) aufgezeichnet. Offline wurde dann für jede Versuchsreihe jeder Einzeltrial (Onset Target/Onset Response) mit den RT-Daten des Testobjektes automatisiert verglichen, bewertet und die durchschnittlichen Abweichungen dokumentiert.

Bei den eingesetzten Computern handelte es sich um folgende Systeme:

PC1 (Standcomputer):

AMD Athlon Barton 2500+, 1833 MHz, ASUS A7N8X2.0,

PC 2 (Standcomputer):

AMD Athlon Barton 2500+, 900 MHz, ASUS A7N8X2.0.

Beide Standcomputer waren mit 2x512 mb RAM der Fa. Corsair bestückt.

PC3 (Notebook):

ASUS A6V, Pentium M740, 2x 512MB RAM, Ati Radeon Mobility X700.

Bei den Monitoren handelte es sich um einen Samsung Syncmaster 1200NF (Röhrenmonitor) mit einer eingestellten Auflösung von 1024 x 768 Bildpunkten. In einer Konfiguration wurde er mit 75 Hz betrieben (1. Monitor), in der anderen Konfiguration wurde er mit 100 Hz betrieben (2. Monitor). Als drit-

ter Monitor wurde ein Viewsonic X924 (TFT-Bildschirm) in das Versuchsdesign aufgenommen, der ebenfalls mit einer Auflösung von 1024 x 768 Bildpunkten während des Versuchs betrieben wurde (das entspricht der vom Hersteller empfohlenen optimalen Auflösung). Die Bildwiederholungsfrequenz betrug 75 Hz.

Die Vorversuche fanden im Dezember 2006/Januar 2007 statt. Im Hauptversuch mit der neueren Version des computergestützten Testsystems zeigte sich, dass die weiterentwickelte Software in einer Windows 98 Umgebung nicht mehr lauffähig war. Das Untersuchungsdesign wurde deshalb reduziert und in den Versuchsreihen zwei statt drei verschiedene Betriebssysteme berücksichtigt. Damit ergeben sich vier Softwarekonfigurationen:

Software 1 („WinXP raw“):

Frische Windows XP Professional Installation, kein ServicePack, nur Netzwerk und Grafiktreiber.

Software 2 („WinXP normal“)⁴:

Windows XP in gängiger Arbeitsumgebung der Software (Windows Update vom 27.02.07).

Software 3 („Win2000 raw“):

Frische Windows 2000 Professional Installation, ServicePack 4, Netzwerk und Grafiktreiber.

Software 4 („Win2000 normal“)⁵:

Windows 2000 in gängiger Umgebung der Software (Windows Update vom 28.02.07).

Bild 1 zeigt das tatsächliche getestete Design noch einmal im schematischen Überblick.

⁴ Die „normale“ WinXP-Installation beinhaltet folgende Softwarekomponenten: Ad-Aware SE Personal, Adobe Flash Player 9, ActiveX, Adobe Photoshop 7.0, Adobe Reader 7.0.7 – Deutsch, Microsoft .NET Framework 2.0, Mozilla Firefox (1.5), Nero 6 Ultra Edition, NVIDIA Drivers, NVMixer, O&O Defrag Professional Edition, OpenOffice.org 2.0, PowerQuest Drive Image 2002, PowerQuest PartitionMagic 8.0, Prime95, Printserver Driver, Quicktime Alternative 1.75, RealAlternative 1.50, Recovery for PowerPoint, SiSoftware Sandra Lite 2007, Sophos Anti-Virus, VideoLAN VLC media player 0.8.5, Windows XP Servicepack 2

⁵ Die „normale“ Windows 2000-Installation beinhaltet folgende Softwarekomponenten: Ad-Aware SE Personal, Adobe Photoshop 7.0, Adobe Reader 7.0.7 – Deutsch, Microsoft .NET Framework 2.0, Microsoft Internet Explorer 6 SP1, Mozilla Firefox (2.0), Nero 6 Ultra Edition, NVIDIA Drivers, NVMixer, O&O Defrag Professional Edition, SiSoftware Sandra Lite 2007, Sophos Anti-Virus, Synchronizable v1.5, VideoLAN VLC media player 0.8.5, Winamp, Windows 2000 Servicepack 4 und WinRAR Archiver.

PC 1 1.833 MHz (11 x 1 66)	WinXP raw
	WinXP normal
	Win2000 raw
	Win2000 normal
PC 2 900 MHz (9 x 100)	WinXP raw
	WinXP normal
	Win2000 raw
	Win2000 normal
Notebook 1.700 MHz (13 x 133)	WinXP raw
	WinXP normal
	Win2000 raw
	Win2000 normal
Monitor 1, CRT, 1024x768, 75 Hz	
Monitor 2, CRT, 1024x768, 100 Hz	
Monitor 3, TFT, 1024x768, 75H z	

Bild 1: Testdesign im schematischen Überblick

10.6 Versuchsdurchführung

Vor jeder Versuchsreihe wurde das System geeicht. Jede Konfiguration wurde in fünf Durchläufen getestet mit jeweils 127 von 128 Items eines normalen Durchlaufes (ein Item konnte wegen der Synchronisation mit dem Rechtecksignal-Generator nicht ausgewertet werden).

Ein „echter“ Proband reagierte auf die auf dem Bildschirm dargebotenen Zielreize und drückte die entsprechende Taste beim mitgelieferten Eingabegerät.

10.7 Ergebnisse

Nachfolgend werden die Ergebnisse für die in Kapitel 10.5 beschriebenen Konfigurationen zusammenfassend dargestellt. Jede Versuchsreihe (Messwertreihe) bestand aus jeweils 127 Einzelmessungen (s. o.) mit jeweils 5 Durchläufen. Die absoluten Reaktionsgeschwindigkeiten werden nachfolgend nicht aufgeführt, da vor allem die gemessenen Abweichungen von Interesse sind (zu den absoluten Reaktionszeiten mehr in Kapitel 10.8).

Insgesamt wurden 36 verschiedene Konfigurationen getestet. Dabei weichen die objektiven Reaktionszeiten von den vom Testverfahren ausgegebenen Reaktionszeiten je nach Konfiguration zwischen maximal 26 ms und minimal 16 ms (Tabelle

5) zu Ungunsten des Probanden voneinander ab.

Innerhalb einer einzelnen Computerkonfiguration streut der Mittelwert zwischen 0 ms und 3 ms bei jeweils 5 Testreihen einer Konfiguration. Die durchschnittliche Abweichung von den objektiven Reaktionszeiten insgesamt – gemittelt über alle Testläufe – beträgt 20,7 ms zu Ungunsten des Probanden, die Standardabweichung für die 180 Messwertreihen beträgt 2 ms. In den nachstehenden Tabellen sind die Häufigkeiten der Abweichungen sortiert nach Millisekunden aufgeführt:

10.7.1 Ergebnisse im Detail

Tabelle 5 zeigt, dass die Konfiguration mit der genauesten Messung im Durchschnitt noch immer eine Abweichung von 16 ms zu Ungunsten des Probanden ermittelt. In keiner Messwertreihe wurde die Reaktionszeit des Probanden genau gemessen. Die ungenaueste Messung trat bei zwei Konfigurationen mit einer durchschnittlichen Abweichung von 26 ms zu Ungunsten des Probanden auf. Deutlich größer sind die maximalen Abweichungen, die bei einzelnen Items gemessen wurden (s. Tab. 6).

In jeder Messwertreihe trat mindestens eine Abweichung von 35 ms oder größer auf. Die maximale Messwertabweichung eines Einzelitems betrug 85 ms.

Die Ausgangsfrage für die Vermessung des getesteten psychometrischen Messverfahrens war, ob es

Mittlere Abweichung in ms	Häufigkeit	Prozent
16 ms	1	,6
17 ms	9	5,0
18 ms	17	9,4
19 ms	26	14,4
20 ms	28	15,6
21 ms	33	18,3
22 ms	29	16,1
23 ms	22	12,2
24 ms	10	5,6
25 ms	3	1,7
26 ms	2	1,1
Gesamt	180	100,0

Tab. 5: Häufigkeit von durchschnittlichen Abweichungen in Millisekunden bei allen Messwertreihen

Max. Abweichung	Häufigkeit	Prozent
31 ms	1	,6
32 ms	2	1,1
33 ms	1	,6
34 ms	5	2,8
35 ms	6	3,3
36 ms	11	6,1
37 ms	6	3,3
38 ms	6	3,3
39 ms	5	2,8
40 ms	8	4,4
41 ms	7	3,9
42 ms	10	5,6
43 ms	6	3,3
44 ms	5	2,8
45 ms	5	2,8
46 ms	10	5,6
47 ms	8	4,4
48 ms	5	2,8
49 ms	9	5,0
50 ms	10	5,6
51 ms	7	3,9
53 ms	4	2,2
54 ms	10	5,6
55 ms	7	3,9
56 ms	3	1,7
57 ms	2	1,1
58 ms	2	1,1
60 ms	1	,6
61 ms	3	1,7
62 ms	3	1,7
63 ms	2	1,1
64 ms	3	1,7
67 ms	2	1,1
70 ms	2	1,1
73 ms	1	,6
76 ms	1	,6
85 ms	1	,6
Gesamt	180	100,0

Tab. 6: Maximale Abweichung eines Einzelitems je Messreihe

Quelle	df	F-Wert	Signifikanz
PC	2	95,745	p < .000
Monitor	2	299,865	p < .000
Software	3	18,133	p < .000
PC * Monitor	4	,259	p < .904
PC * Software	6	7,161	p < .000
Monitor * Software	6	,699	p < .651
PC * Monitor * Software	12	2,766	p < .002

Tab. 7: Varianzanalyse des Einflusses der Systemkonfiguration auf die Messwerte

in Abhängigkeit der Systemkonfiguration zu unterschiedlichen Messergebnissen kommen kann. Hierzu wurde über die 180 Messwertreihen eine Varianzanalyse durchgeführt bei der angenommen wurde, dass es sich bei den Einzelmessungen um voneinander unabhängige Messungen handelt. Daraus ergibt sich ein 3x3x4-Design für die Varianzanalyse.

Im Ergebnis kann festgehalten werden, dass alle Ursprungsvariablen (Computer, Monitor und Softwarekonfiguration) einen überzufälligen Einfluss auf das Messergebnis haben. In Tabelle 7 sind die Ergebnisse der Varianzanalyse zusammengefasst.

10.7.2 Einfluss der Einzelkomponenten

Werden die Reaktionszeiten nur in Abhängigkeit des Computersystems dargestellt, ergeben sich folgende Mittelwertunterschiede (siehe Tabelle 8).

Wie die Tabelle 8 zeigt, gibt es signifikante Mittelwertunterschiede zwischen den Computersystemen, die sich im Bereich von insgesamt 2 ms bewegen.

Werden die Mittelwertunterschiede nur in Abhängigkeit der Monitorkonfiguration dargestellt, ergeben sich die in Tabelle 9 dargestellten Werte.

Die erfassten Mittelwerte der Reaktionsgeschwindigkeiten unterscheiden sich in Abhängigkeit vom Monitor wiederum überzufällig (s. Tabelle 9) und bewegen sich hier im Bereich von bis zu etwa 4 ms. Bezogen auf die Einzelkomponenten des Messsystems hat damit der Monitor den größten Einfluss auf Mittelwertunterschiede. Wie Tabelle 10 zeigt, bewegen sich die gemessenen Mittelwertunterschiede in Abhängigkeit der eingesetzten Software in einem sehr engen Bereich zwischen 20,1 ms und 21,2 ms.

PC	Mittelwert	N	Standardabweichung
PC 1, 1.833 Mhz	21,7 ms	60	2,0
PC 2, 900 Mhz	21,0 ms	60	1,8
Notebook, 1.700 Mhz	19,6 ms	60	1,8
Insgesamt	20,7 ms	180	2,0

Tab. 8: Mittelwertunterschiede in Abhängigkeit der Computerkonfiguration

PC	Mittelwert	N	Standardabweichung
Mon. 1, Röhre, 100 Hz	20,8 ms	60	1,4
Mon. 2, Röhre, 75 Hz	18,9 ms	60	1,4
Mon. 3, TFT, 75 Hz	22,6 ms	60	1,3
Insgesamt	20,7 ms	180	2,0

Tab. 9: Mittelwertunterschiede in Abhängigkeit der Monitorkonfiguration

Software	Mittelwert	N	Standardabweichung
XP raw	20,5 ms	45	1,8
XP norm	21,1 ms	45	2,2
W2K raw	20,1 ms	45	1,7
W2K norm	21,2 ms	45	2,3
Insgesamt	20,7 ms	180	2,0

Tab. 10: Mittelwertunterschiede in Abhängigkeit der Softwarekonfiguration

Obwohl auch hier die gemessenen Mittelwertunterschiede überzufällig sind (s. Tabelle 8), schwanken sie im Bereich von 1 ms.

10.7.3 Wechselwirkungen der Komponenten untereinander

Wie Tabelle 7 zeigt, gibt es neben den Wirkungen der Hauptfaktoren auch noch Wechselwirkungen der Komponenten untereinander, die ebenfalls überzufällige Unterschiede in den Messwertreihen zeigen. Nicht überzufällig sind die Wechselwirkung zwischen PC und Monitor und die Wechselwirkung zwischen Monitor und Software.

Einen überzufälligen Zusammenhang zu den Reaktionszeitmessungen zeigen die Wechselwirkung zwischen PC und Software und die „Dreifachwechselwirkung“ aller Komponenten untereinander. Zur Verdeutlichung werden in Tabelle 11 alle Konfigurationen – geordnet nach Abweichungen zu den tatsächlichen Reaktionszeiten – dargestellt.

Konfigurations- Nr.	Mittelwert	N	Standardabweichung
35	17,2 ms	5	0,4
29	17,4 ms	5	1,1
26	18,0 ms	5	1,2
17	18,2 ms	5	0,8
32	18,2 ms	5	0,4
2	18,6 ms	5	0,5
31	18,8 ms	5	0,4
14	19,0 ms	5	0,7
28	19,0 ms	5	0,7
5	19,2 ms	5	0,8
11	19,6 ms	5	1,1
23	19,6 ms	5	0,9
34	20,0 ms	5	0,7
25	20,2 ms	5	0,8
1	20,6 ms	5	0,9
4	20,6 ms	5	0,5
8	20,6 ms	5	0,5
16	20,6 ms	5	0,5
13	20,8 ms	5	0,8
20	20,8 ms	5	1,3
33	21,0 ms	5	1,0
10	21,4 ms	5	0,5
22	21,4 ms	5	0,5
36	21,4 ms	5	1,1
27	21,6 ms	5	0,9
18	22,0 ms	5	0,0
19	22,0 ms	5	0,7
30	22,2 ms	5	0,4
15	22,4 ms	5	0,9
6	22,6 ms	5	0,5
3	23,0 ms	5	0,7
12	23,0 ms	5	0,0
7	23,6 ms	5	0,5
9	23,6 ms	5	0,9
24	23,6 ms	5	1,9
21	24,8 ms	5	0,8
Insgesamt	20,7 ms	180	2,0

Tab. 11: Mittelwertunterschiede bei den Reaktionszeiten in Abhängigkeit der Gesamtkonfiguration

Wie Tabelle 11 zeigt, streuen die Mittelwerte in Abhängigkeit der jeweiligen Gesamtkonfiguration von etwa 17 ms bis zu 25 ms. Das bedeutet eine Streuung von etwa 4 ms um den Gesamtmittelwert von 20,7 ms in jede Richtung. Hier wird besonders deutlich, dass nicht die Streuung um den Gesamtmittelwert das Hauptproblem darstellt, sondern die grundsätzliche Eichung des Testsystems an sich (vgl. auch ULRICH & GIRAY, 1989; WARTENBURG, F. (ohne Jahresangabe)). Es gibt zwar in Abhängigkeit der Konfiguration statistisch überzufällige Unterschiede von etwa 4 ms in jede Richtung um den Mittelwert, ob sie jedoch auch für die Prüfung der Fahreignung im Leistungstest bedeutsam sind, kann an dieser Stelle nicht entschieden werden. Es kann aber davon ausgegangen werden, dass Streuungen in dieser Größenordnung zu vertreten sind. Abschließend könnte diese Frage nur anhand von Normwerttabellen entschieden werden, die jedoch nicht aus dem untersuchten Testsystem direkt ausgelesen werden konnten und auch in gedruckter Form nicht vorlagen. Deshalb musste in einer zweiten Experimentalreihe ermittelt werden, welche Normwerte den Messungen zu Grunde liegen, um entscheiden zu können, ob die ermittelten Reaktionszeitunterschiede in Abhängigkeit der Systemkonfiguration auch für den Probanden bedeutsam sein können, wenn seine Leistungen in die Prozentrangwerttabelle eingeordnet werden.

10.8 Ermittlung der Normwerttabellen zur Überprüfung der Bedeutsamkeit der festgestellten Mittelwertabweichungen

Alleine aus den Abweichungen zwischen objektiven Reaktionszeiten und den vom untersuchten Testsystem ermittelten Reaktionszeiten lässt sich nicht abschätzen, welche Auswirkungen die Ungenauigkeiten bei der Einordnung in Prozentrangwerte beim Reaktionszeittest haben. Ob die ermittelten Ungenauigkeiten von durchschnittlich 21 ms zu Ungunsten eines getesteten Probanden wirklich dazu führen könnten, den Reaktionszeittest nicht zu bestehen, ist allein aus der Höhe der Abweichungen nicht zu entscheiden. Vielmehr muss die Höhe der Abweichungen in Relation zu den dem Testverfahren zu Grunde liegenden Normwerttabellen gesehen werden.

Da aus der Software des untersuchten Testverfahrens keine Normwerttabellen ausgelesen werden konnten und sich auch im Handbuch keine Norm-

werttabellen finden ließen, in denen die Prozentrangverteilung für die Reaktionszeitmessungen dargestellt waren, musste auf andere Art versucht werden zu ermitteln, wie sehr sich die nachgewiesenen Mittelwertunterschiede zu den tatsächlichen Messwerten bei einer „echten“ Leistungsprüfung auswirken könnten.

Es muss noch einmal erwähnt werden, dass sich die gemessenen Mittelwertunterschiede von den tatsächlichen Mittelwerten in der Größenordnung von durchschnittlich 21 ms alle zu Ungunsten eines möglichen Probanden gezeigt haben.

10.8.1 Versuchsbeschreibung zur Ermittlung der dem untersuchten Testverfahren zu Grunde liegenden Normwerttabellen

Da es zur Ermittlung der Normwerttabellen nötig war, die Antwortreaktionszeiten in sensiblen Bereichen (16. und 33. Prozentrang) in 1-ms-Schritten zu verändern, kam es nicht in Frage, dieses Experiment wieder mit einem echten Probanden durchzuführen. Vielmehr wurde nun ein „elektronischer Proband“ eingesetzt, bei dem das „Antwortverhalten“ für die in Frage kommenden sensiblen Rangwerte (16. und 33. Perzentil) in 1-ms-Schritten experimentell variiert werden konnte. Diese Testreihe wurde jedoch mit nur einer Testkonfiguration durchgeführt und für jede Veränderung der Reaktionszeit mit nur einer Testreihe.

Hard-/Softwarekonfiguration

Stand-PC, 1833MHz; neue Windows 2000 Professional-Installation, ServicePack 4, Netzwerk und Grafiktreiber; Monitor Samsung Syncmaster 1200NF, Auflösung 1024x768 bei 75Hz.

Das Messsystem wurde vor Durchführung der Messwertreihen geeicht und während der Messungen noch einmal bei den eingestellten Zeiten von 600 ms, 700 ms und 800 ms.

10.8.2 Ergebnisse zur Ermittlung der dem Test zu Grunde liegenden Normwerttabellen

In Tabelle 12 sind die Ergebnisse dieser Testmessungen dargestellt. In der ersten Spalte ist die geeichte Reaktionszeit des elektronischen Probanden eingetragen. In der zweiten Spalte ist der vom untersuchten Testsystem ermittelte Testwert eingetragen, in der dritten Spalte die gemessene Differenz und in der vierten Spalte der vom untersuchten

Reaktionszeit des elektronischen Probanden in ms	Vom Test ausgegebener Messwert in ms	Differenz in ms	Vom Test ausgegebener Prozentrang
480	503,2	23,2	100
490	514,2	24,2	95
500	524,4	24,4	95
510	534,1	24,1	95
520	545,0	25,0	84
530	554,9	24,9	84
540	562,6	22,6	84
550	572,8	22,8	79
560	583,3	23,3	76
570	593,3	23,3	76
580	603,7	23,7	69
590	613,6	23,6	66
595	619,6	24,6	66
600	625,7	25,7	62
605	627,5	22,5	62
610	635,2	25,2	54
615	639,4	24,4	54
620	645,2	25,2	50
625	649,9	24,9	50
630	654,3	24,3	46
640	663,4	23,4	46
645	668,8	23,8	38
650	673,9	23,9	38
651	674,6	23,6	38
652	674,5	22,5	38
653	676,6	23,6	34
654	678,8	24,8	34
655	679,3	24,3	34
656	679,6	23,6	34
657	682,3	25,3	34
658	681,6	23,6	34
659	683,9	24,9	31
660	684,4	24,4	31
670	694,8	24,8	31
675	700,1	25,1	24
680	705,6	25,6	24
685	708,7	23,7	21
690	714,4	24,4	21
695	719,4	24,4	21
696	719,9	23,9	21
697	723,0	26,0	16
698	723,3	25,3	16
699	723,3	24,3	16
700	726,3	26,3	16
...			

Tab. 12: Vergleich zwischen objektiven und vom Testsystem ermittelten Reaktionszeiten

Reaktionszeit des elektronischen Probanden in ms	Vom Test ausgegebener Messwert in ms	Differenz in ms	Vom Test ausgegebener Prozentrang
710	735,0	25,0	16
711	736,4	25,4	16
712	737,9	25,9	16
713	738,3	25,3	10
714	740,9	26,9	10
715	739,6	24,6	10
720	744,4	24,4	10
725	751,8	26,8	10
730	754,6	24,6	5
740	763,7	23,7	5
750	774,1	24,1	5
760	785,5	25,5	5
770	794,4	24,4	5
780	804,9	24,9	5
790	814,2	24,2	5
800	825,7	25,7	0

Tab. 12: Fortsetzung

Testsystem ausgegebene Prozentrangwert. Grau unterlegt sind Prozentrangwertbereiche, die für die Fahrerlaubnis eine Rolle spielen. Dort wurden die Reaktionszeiten in 1-ms-Schritten variiert. Auslassungen in der Darstellung sind mit „...“ gekennzeichnet, wenn sich in dem Bereich trotz Änderung der voreingestellten Reaktionszeiten keine Änderungen bei der Zuordnung zu Prozentwertungen gezeigt haben.

Insgesamt zeigt sich, dass bei der Messreihe mit einem „elektronischen Probanden“ die durchschnittlichen Abweichungen vom objektiven Messwert noch ein wenig größer sind als bei dem Test mit einem echten Probanden und dass Fehlmessungen in für die Fahreignung sensiblen Bereichen sehr wohl eine deutliche Verschiebung bei der Einordnung in die Prozentrangwerte zu Ungunsten des Probanden bedeuten können.

Einer der auffälligsten Befunde bei der Auswertung der Tabelle 12 ist, dass auch in sensiblen Bereichen, bei denen darüber entschieden wird, ob die Voraussetzungen nach FeV, Anlage 5, überhaupt vorliegen, keine diskreten Einzel-Prozentrangwerte vom System ausgegeben werden. So bedeutet ein Sprung der Reaktionszeit um 1 ms im Bereich von 713 ms zwar einen Sprung von Rangplatz 10 auf Rangplatz 16 (Anforderung erfüllt), aber Rangplatz 16 wird auch beibehalten, wenn die Antwort 15 ms schneller erfolgt (im Bereich 697 ms). Untersucher,

die darüber entscheiden müssen, ob jemand die Mindestanforderungen bei den Reaktionszeiten erfüllt oder nicht, können also gar nicht einschätzen, in welchem Bereich sich die Leistung des Probanden genau befindet. Ähnliches gilt für den Bereich vom 33. Perzentil, der vor allem für Fahrer der Gruppe 2 eine Rolle spielt. Auch hier werden keine diskreten Einzel-Rangwerte vom untersuchten Testsystem ausgegeben.

Beispiel

Angenommen, der Proband reagiert in Wirklichkeit mit einer Reaktionszeit von 730 ms (Spalte: „Reaktionszeit des elektronischen Probanden in ms“), dann müsste er bei den vom System ausgegebenen Prozentrangwerten eigentlich in die Kategorie Prozentrangwert 16 (Anforderung erfüllt) fallen (siehe Spalte „Vom Test ausgegebener Messwert in ms“). Wegen der Fehlmessung werden vom System aber 754,6 ms gemessen und der Proband fällt in Gruppe Prozentrangwert 5. Der Untersucher muss also davon ausgehen, dass der vermessene Proband die Anforderungen bei weitem nicht erfüllt hat, obwohl er objektiv schnell genug reagiert hat. Dabei gilt, dass besonders bei eher leistungsschwächeren Probanden eine erhebliche Fehleinschätzung der gezeigten Leistung im Bereich Reaktionszeitmessung auftreten kann. Verschärft wird das Problem dadurch, dass keine diskreten Einzel-Prozentrangwerte nachgelesen werden können und die automatisierte Auswertung des getesteten Systems diese Informationen nicht bereithält. Das Untersuchungspersonal ist hierdurch nicht in der Lage, in den kritischen Bereichen die Leistung des Probanden gut einschätzen zu können, etwa, ob eine Leistung „gerade eben nicht erbracht wurde“ oder „ob der Proband wirklich weit unter den Mindestanforderungen“ im Bereich der Reaktionszeitmessung liegt. Wie das Beispiel am Anfang des Absatzes zeigt, kann es hierdurch bei bestimmten Konstellationen zu erheblichen Fehleinschätzungen kommen.

10.9 Abschließende Beurteilung des Einflusses unterschiedlicher Computerkonfigurationen auf die Reaktionszeitmessung

Die Ursprungsfrage des Erweiterungsprojektes, ob bestimmte Systemkonfigurationen dazu führen können, dass die gemessenen Reaktionszeiten al-

lein durch unterschiedliche Konfigurationen substantiell verändert werden, muss mit zwei Antworten beantwortet werden: Zum einen muss über die grundsätzliche Messgenauigkeit eines Testsystems nachgedacht werden, zum anderen über die Streuung der ermittelten Messwerte, die sich durch unterschiedliche Computerkonfigurationen ergeben kann.

10.9.1 Grundsätzliche Eichung des Systems

Grundsätzlich ist festzuhalten, dass die Basis-Eichung des untersuchten Testsystems zur Reaktionszeitmessung im Rahmen der FeV, Anlage 5, nicht gegeben war. Eine durchschnittliche Fehlmessung von knapp 21 ms zu Ungunsten der Probanden kann besonders bei eher leistungsschwächeren Probanden dazu führen, dass die Leistung im Bereich „Reaktionsfähigkeit“ durch einen Untersucher erheblich falsch eingeschätzt werden kann, obwohl der Proband die geforderte Mindestleistung (16. Perzentil, vgl. Kapitel 10.8.2) im Reaktionszeittest eigentlich erbracht hat. Gravierend ist hier die grundsätzlich fehlende bzw. falsche Eichung des Basissystems. Dieser grundsätzliche Messwertfehler wurde auch bei der „genauesten“ Messung nicht ausgeglichen und betrug noch immer 16 ms zu Ungunsten eines Probanden. Grundsätzlich muss gewährleistet sein, dass Messsysteme im Auslieferungszustand so geeicht sind, dass sie in den Computerumgebungen, für die sie vorgesehen sind, im Durchschnitt genau messen. Bei einer durchschnittlichen Fehlmessung von knapp 21 ms kann von einer zuverlässigen Basiseichung des Messsystems nicht ausgegangen werden. Es ist daher anzustreben (und zu fordern), dass ausgelieferte Testsysteme (wie bei anderen computergestützten Testverfahren auch) in den Computerumgebungen, für die sie programmiert und vom Hersteller zugelassen sind, im Durchschnitt bei unterschiedlichen Konfigurationen keinen Messfehler ausweisen. Der durchschnittliche zu erwartende Messfehler muss „0“ sein.

Das Hauptproblem bei den nachgewiesenen Messfehlern ist eindeutig die grundsätzlich ungenaue Eichung des Messsystems. Sie ist für die Fehleinschätzung der Leistung von Probanden in Reaktionszeittests hauptverantwortlich.

10.9.2 Einfluss unterschiedlicher Hard-/und Softwarekonstellationen auf die Messung

Wird nur der Einfluss der Computerkonfiguration auf den Messwertfehler betrachtet, zeigt die Untersuchung, dass je nach Computerkonfiguration der Messfehler im Mittel um etwa -4 ms bis etwa +4 ms um den Mittelwert von 21 ms Fehlmessung streut. Die gemessenen Abweichungen, abhängig von den unterschiedlichen Computerkonfigurationen, sind zwar auch statistisch überzufällig (Gesamtstreuung aller Testreihen zwischen +16 ms und +26 ms bei einzelnen Testreihen), trotzdem dürften Fehleinschätzungen – in der Größenordnung von max. 5 ms im Einzelfall um den Mittelwert – zu vertreten sein. Allerdings muss dann die Möglichkeit gegeben sein, die Abweichungen in Normwerttabellen nachzulesen, um eine angemessene Einschätzung der vom Probanden gezeigten Leistung ermöglichen zu können (siehe Kapitel 10.9.3). Wäre das computergestützte Testsystem insgesamt besser geeicht und würden die konfigurationsbedingten Abweichungen insgesamt um einen Mittelwert von 0 ms Fehlmessungen streuen, wäre die Abweichung für einzelne Konfigurationen trotz überzufälliger Unterschiede zu vertreten. Allerdings müssten/müssen Hersteller von computergestützten Tests im Rahmen der Fahreignungsprüfung diese maximal möglichen Abweichungen für die zugelassenen Konfigurationen angeben und/oder geeichte Systemkonfigurationen mit eben diesen Angaben vorschreiben.

10.9.3 Angabe von Prozentrangwerten und Konfidenzintervallen für die Auswertung

Unabhängig vom Messwertfehler ist es für jeden Untersucher unerlässlich, bei einem psychologischen Leistungstest, der auch für die Fahrerlaubnis relevant ist, die Leistung eines Probanden in den richtigen Kontext stellen zu können. Da die Begutachtungs-Leitlinien Prozentrangwerte vorschreiben und auch die Möglichkeit besteht, Leistungen, die den Anforderungen nicht entsprechen, durch Leistungen in anderen Bereichen zu kompensieren, sollte der Untersucher in der Lage sein können, die gemessenen Leistungen sehr genau in Relation zur Normstichprobe setzen zu können. Dazu ist es unerlässlich, diskrete Normwerte für jeden Prozentrangwert nachlesen zu können, ganz besonders jedoch in den für die Fahreignung sensiblen Bereichen (16. und 33. Perzentil).

Wenn für das computergestützte Testsystem der maximal mögliche Fehler angegeben und die ermittelte Leistung vom Untersucher begutachtet wird, sollte er nachlesen können, ob die erbrachte Leistung wirklich in einem Bereich liegt, der nicht den Anforderungen genügt oder durch einen möglichen Messfehler ein Wert erreicht wurde, der den Anforderungen möglicherweise genügt und lediglich messfehlerbehaftet ist.

Hierzu sind zwei Voraussetzungen nötig:

Zum einen muss der maximal mögliche Messfehler bekannt sein und vom Hersteller angegeben werden (s. o.). Bei der automatisierten Auswertung und Ausgabe von Berichten zur psychometrischen Leistungsfähigkeit sollte, wie es etwa auch bei Intelligenztests oder anderen psychologischen Messverfahren üblich ist, angegeben werden, mit welcher Wahrscheinlichkeit der gemessene Leistungswert in welchem Intervall liegt (Konfidenzintervall). Dieses Intervall sollte für jeden Prozentrangwert angegeben werden.

Um eine gemessene Leistung darüber hinaus in den richtigen Kontext stellen zu können, muss es zusätzlich möglich sein, die vollständigen Normwerttabellen im Überblick einsehen zu können. Nur hierdurch kann ein Untersucher entscheiden, ob eine Leistung eher am unteren Rand oder eher am oberen Rand einer bestimmten Leistungskategorie einzuordnen ist.

11 Zusammenfassung und Ausblick

In dem Forschungsprojekt „Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung“ (FE 82.291/2005) wurden die gängigen in Deutschland eingesetzten Verfahren zur psychometrischen Leistungsprüfung der Fahreignung dahingehend analysiert, ob sie – wie es die FeV, Anlage 5 vorschreibt – auf der Basis wissenschaftlicher Kriterien standardisiert sind und unter Aspekten der Verkehrssicherheit validiert wurden.

Außerdem wurde in einem zweiten Projektteil der Frage nachgegangen, ob die Nutzung unterschiedlicher PC-Systeme, Monitore und Softwarekonfigurationen einen Einfluss auf das vom Computertest ermittelte psychologische Leistungsmerkmal „Reaktionsfähigkeit“ hat.

Schon bei der ersten Durchsicht der wissenschaftlichen Literatur zu diesem Thema zeigte sich, dass die in der FeV genannten und mit Hilfe der Testverfahren zu überprüfenden Bereiche „Belastbarkeit“, „Orientierungsleistung“, „Konzentrationsleistung“, „Aufmerksamkeitsleistung“ und „Reaktionsfähigkeit“ neueren wissenschaftlichen Taxonomien nicht mehr entsprechen.

Um Aussagen dazu machen zu können, wie Begutachtungsstellen mit dieser Problematik umgehen, welche Probleme in der Praxis beim Einsatz der gängigen psychometrischen Testverfahren auftauchen, welche Schwierigkeiten bei der Anwendung der computergestützten Testverfahren bestehen und welche Verbesserungsmöglichkeiten in der Praxis gesehen werden, wurden fast alle bundesdeutschen Begutachtungsstellen zu dieser Problematik angeschrieben. Die grundsätzliche Idee bestand darin, mit Unterstützung der Begutachtungsstellen ein mögliches Verbesserungspotenzial zu eruieren und, unterstützt durch die Erfahrungen aus der Praxis, Vorschläge zu einer Verbesserung der gegenwärtigen Situation und einer inhaltlichen Schärfung der psychometrischen Leistungsbereiche zu erarbeiten. Die Befragung wurde von den Begutachtungsstellen allerdings weitgehend nicht beantwortet, sodass in dem Forschungsprojekt auf die als wertvoll erachteten Beiträge aus der Praxis beim Erkenntnisgewinn verzichtet werden musste. Insofern wurde das Forschungsprojekt noch stärker theoriegeleitet durchgeführt als ursprünglich geplant.

Kernstück der Untersuchung bildet die kritische Durchsicht der in der Praxis eingesetzten Verfahren zur psychometrischen Leistungsprüfung der Fahreignung anhand einer wissenschaftlichen Kriterienliste und die anschließende Diskussion der Studienergebnisse bei einem eintägigen Workshop mit Experten aus der Wissenschaft und Praxis.

Grundlage für die inhaltliche Entwicklung der Kriterienliste waren die für die eingesetzten Verfahren in der Fahrerlaubnisverordnung geforderte wissenschaftliche Standardisierung und Validierung unter Aspekten der Verkehrssicherheit. Die Kriterienliste für die kritische Durchsicht wurde auf Basis gängiger wissenschaftlicher Standardwerke zur Konstruktion von Testverfahren entwickelt und eng mit der Bundesanstalt für Straßenwesen abgestimmt. Durch die einheitliche Kriterienliste wurde gewährleistet, alle Testverfahren nach den genau gleichen Kriterien einzuschätzen und die Ergebnisse ver-

gleichbar zu machen. Die Kriterienliste bezieht sich vor allem auf einfache Testgütekriterien, die jedem wissenschaftlichen Testverfahren zu Grunde liegen sollten, aber auch auf die theoretisch wissenschaftliche Fundierung der Testverfahren. Ein weiterer Fokus lag auf der Fragestellung, ob die Testverfahren auch an der Population normiert wurden, für die sie eingesetzt werden, und der Dokumentation der Testentwicklung sowie der Testgütekriterien. Wie ein Test entstanden ist, welche theoretischen Konstrukte den gemessenen psychologischen Konstrukten zu Grunde liegen, wie die in den Manualen beschriebenen Testgütekriterien ermittelt wurden und welche Normwerttabellen wie konstruiert wurden, sollte für die Anwender in der Praxis nachvollziehbar und gut dokumentiert sein. Nur so kann wirklich entschieden werden, ob ein computergestütztes Testverfahren zur psychometrischen Leistungsprüfung auch tatsächlich die Bereiche misst, die es vorgibt zu messen bzw. die es messen soll.

Im Ergebnis zeigt sich, dass bei den derzeit eingesetzten computergestützten psychometrischen Testverfahren zur Fahreignung noch einiger Optimierungsbedarf besteht. Besonderer Optimierungsbedarf wird bei der Normierung der Testverfahren gesehen und der nachvollziehbaren Dokumentation der Testgütekriterien in den Manualen. Die von den Testautoren berichteten Testgütekriterien sind nicht immer nachvollziehbar; in Einzelfällen fehlen sogar fundamentale Angaben hierzu. Auch werden nicht in jedem Einzelfall die Einzeltests der Testbatterien den in der FeV genannten psychometrischen Leistungsbereichen eindeutig zugeordnet. Die wissenschaftlich theoretischen Grundlagen der zu messenden Konstrukte sind ebenfalls oft unzureichend dokumentiert.

Nach der Analyse der Testverfahren wurden in einem Expertenworkshop die ermittelten Projektzweischenergebnisse diskutiert. In dem Workshop wurden den von der BASt eingeladenen Experten aus den Bereichen Methodenlehre, Testkonstruktion, Eignungsdiagnostik und Neuropsychologie das gewählte Vorgehen und die vorläufigen Ergebnisse des Forschungsprojektes vorgestellt. Insgesamt zeigte sich, dass von den Experten das gewählte Vorgehen als wissenschaftlich angemessen beurteilt wurde und die ermittelten Ergebnisse dem derzeitigen wissenschaftlichen Erkenntnisstand entsprechen. Auch die erarbeitete Kriterienliste wurde als angemessen beurteilt. Durch die Expertise der Experten sollten die bis zu diesem Zeitpunkt

ermittelten Forschungsergebnisse noch einmal zusätzlich wissenschaftlich abgesichert werden. Inhaltlich gab es faktisch keine Differenzen zwischen den ermittelten Forschungsergebnissen und der durch die Experten vertretenen Meinungen bezüglich der wesentlichen Optimierungsvorschläge für die computergestützten Testverfahren.

In der Diskussion wurde bemängelt, dass es bisher im Gegensatz zu anderen Bereichen im Kontext der Fahreignung keine festgelegten Standards gibt, die die eingesetzten Testverfahren verbindlich einhalten müssen.

Beim Expertentreffen wurde deutlich, dass in den Manualen der Testverfahren auf jeden Fall alle basalen Testgütekriterien aufgeführt und gut nachvollziehbar dokumentiert sein müssen. Besonderer Wert wurde in der Diskussion auf die theoriebasierte Konstruktvalidität, die Kriteriumsvalidität und die Angabe der prognostischen Validität gelegt. Die Hauptergebnisse des Expertentreffens beziehen sich auf eine kurzfristige und eine langfristige Perspektive zum Einsatz psychometrischer Verfahren in der Fahreignungsdiagnostik. Diese Ergebnisse stellen auch gleichzeitig die Hauptergebnisse des vorliegenden Forschungsprojektes dar:

Kurzfristig

Für die derzeit eingesetzten Testverfahren besteht bei allen Verfahren aus Sicht der Experten noch ein mehr oder weniger großer Optimierungsbedarf. Alle Testverfahren sollten anhand von vorher festgelegten Kriterien überarbeitet werden. Als Basis könnte die im Forschungsprojekt entwickelte Kriterienliste dienen.

Im Einzelnen sollten die Testverfahren an ausreichend großen repräsentativen Stichproben geeicht werden, für die sie später auch zugelassen sind (alle potenziellen Fahrer der Gruppe 1, alle potenziellen Fahrer der Gruppe 2). Die Testgütekriterien für den jeweiligen Einzeltest müssen nachvollziehbar angegeben werden, auch auf welcher Datenbasis sie im Einzelfall ermittelt wurden. Das sind im Wesentlichen genaue Angaben zur Eichstichprobe, zu den Reliabilitätsmaßen und den zu Grunde liegenden Stichproben, zu Stabilitätsmaßen (prognostische Güte auch für die Einzeltests) sowie Angaben zur Konstrukt- und Kriteriumsvalidität, die wissenschaftlich hergeleitet sind. Die Einzeltests der Testbatterien müssen sich den in der FeV genannten Bereichen eindeutig zuordnen lassen.

Nachdem die Tests optimiert wurden, sollte eine noch einzurichtende unabhängige wissenschaftliche Instanz die Einhaltung der vorher festgelegten Kriterien überprüfen. Besonderer Wert wurde auf die Unabhängigkeit der Gutachter gelegt, auch darauf, dass es mehr als ein Gutachter sein soll, der die Einhaltung der festgelegten Kriterien überprüft. Erst wenn diese Voraussetzungen erfüllt sind, sollte ein psychometrisches Testverfahren zur Überprüfung der Fahreignung (weiterhin) zugelassen werden dürfen.

Langfristig

Langfristig wird nicht nur ein Optimierungsbedarf für die derzeit eingesetzten Testverfahren gesehen, sondern auch für die gesetzlichen Rahmenbedingungen. Längerfristig müssen aus wissenschaftlicher Perspektive die in der FeV, Anlage 5, genannten wissenschaftlichen Konstrukte (Belastbarkeit, Orientierungsleistung, Konzentrationsleistung, Aufmerksamkeitsleistung, Reaktionsfähigkeit) grundsätzlich überarbeitet werden, da sie nicht dem Stand der aktuellen Forschung entsprechen und somit auch nur unscharf operationalisierbar sind. Darauf aufbauend sollen Tests zur Überprüfung der neueren wissenschaftlichen Konstrukte entwickelt oder eingesetzt werden, sofern sie schon existieren. Aus dieser Zielsetzung ergibt sich ein erheblicher Forschungsbedarf. Auch langfristig sollte ein unabhängiges wissenschaftliches Gremium die Güte der zugelassenen Testverfahren regelmäßig überprüfen. Eine theoretische Fundierung der Testentwicklung für psychometrische Testverfahren zur Überprüfung der Fahreignung wird übereinstimmend als dringend erforderlich angesehen. Dabei sollte kriteriumsorientiert vorgegangen werden: Beginnend mit einer Anforderungsanalyse des Fahrverhaltens im Realverkehr ist festzustellen, welche kognitiven Fähigkeiten mindestens vorliegen müssen, um die Fahraufgabe zu erfüllen, und anschließend die passenden Tests zu suchen bzw. zu entwickeln. Langfristig sollten die in der Fahreignungsdiagnostik angewendeten Entscheidungskriterien (Prozentränge) durch „echte Kriterien“ ersetzt werden, also durch absolute Kennwerte, die auf dem Wissen über die Zusammenhänge zwischen Leistung und Fahreignung basieren. Diese Zusammenhänge sind noch nicht ausreichend erforscht.

Im Anschluss an den Expertenworkshop wurde noch einmal in der wissenschaftlichen Literatur recherchiert, welche Testverfahren zur psychometri-

schen Leistungsprüfung eng mit dem Kriterium „Fahrprobe“ zusammenhängen. Insgesamt existiert zu dieser Fragestellung jedoch nur eine recht spärliche Datenlage im Hinblick auf die im deutschsprachigen Raum eingesetzten Verfahren. Dabei zeigte sich zunächst, dass die in Deutschland gängigen Verfahren wie ART 2020, Expertensystem Verkehr (bzw. Wiener Determinationsgerät und Wiener Reaktionsgerät) und TAP-M zumindest eine gewisse Vorhersage der Fahreignung erlauben. Gleichwohl ist festzustellen, dass die prädiktive Validität der psychometrischen Tests in den genannten Studien sich als allenfalls mittelmäßig erweist, was sowohl auf mangelnde Reliabilität oder Validität der Testverfahren, jedoch auch durchaus auf die Art der Messung des Kriteriums zurückgeführt werden kann. Für den Corporal-A konnten allerdings keine Forschungsergebnisse gefunden werden, bei denen die Testleistungen anhand einer Fahrprobe validiert wurden.

Wenn langfristig bei den Testverfahren stärker kriteriumsorientiert vorgegangen werden sollte, müsste das Außenkriterium „Fahreignung“ bzw. „Fahrprobe“ besser als bisher operationalisiert werden. Bei Durchsicht der Literatur fiel auf, dass das Außenkriterium der Fahrprobe sehr unterschiedlich realisiert wird. Dies wirft zum einen das Problem auf, dass unterschiedliche Studien mit verschiedenen durchgeführten Fahrproben schwer vergleichbar sind. Zum anderen ist fraglich, ob jede Art der Durchführung einer Fahrprobe als Kriterium ausreichend ist. Am sinnvollsten erscheint es, langfristig mittels der psychometrischen Leistungstests die Fahreignung auf folgende Weise vorherzusagen: Die Angabe von Konfidenzintervallen oder wahlweise eine Abstufung von Cut-offs im Sinne von „fahrgeeignet“/„möglicherweise fahrgeeignet: Fahrprobe erforderlich“/„nicht fahrgeeignet“ wäre in der Praxis die hilfreichste Methode. So könnte dem Testanwender eine Hilfestellung gegeben werden, bei welchen Probanden eine nachgeschaltete Fahrprobe nötig ist und welche Probanden sicher klassifiziert werden können.

Bei den Schlussbetrachtungen zu neueren aufmerksamkeits-theoretischen Aspekten im Hinblick auf die in der FeV genannten Bereiche zeigt sich, dass die bisherigen Konstrukte und mögliche zukünftig in der FeV verankerte Konstrukte zur psychologischen Leistungsfähigkeit gar nicht so weit auseinanderliegen müssen. Als möglicher Ausgangspunkt für die zukünftig zu testenden, theoriegeleiteten Bereiche der Aufmerksamkeit werden

zunächst folgende Konstrukte der neueren Aufmerksamkeitsforschung vorgeschlagen:

Intensitätsaspekt

- Allgemeine Reaktionsbereitschaft (Alertness; entspricht weitgehend dem aktuellen Bereich „Reaktionsfähigkeit“),
- Daueraufmerksamkeit (im Sinne von „Belastbarkeit“).

Selektive, räumliche Aufmerksamkeit

- Räumliche Aufmerksamkeitsausrichtung (ähnlich dem jetzigen Bereich „Orientierung“).

Selektive nicht-räumliche Aufmerksamkeit (exekutive Aufmerksamkeit)

- Aufmerksamkeitsteilung (entspricht teilweise dem jetzigen Bereich „Aufmerksamkeit“),
- Fokussierte Aufmerksamkeit (entspricht teilweise dem Bereich „Konzentration“ und beinhaltet die Fähigkeit, sich nicht durch störende visuelle und/oder auditive Reize ablenken zu lassen),
- Strategie/Flexibilität.

Dabei wäre es wichtig, in Zukunft präzisere Operationalisierungen dieser neu formulierten Anforderungsbereiche anzugeben und vor allem auf spezifische Situationen im Straßenverkehr zu beziehen.

Im Hinblick auf neu zu konstruierende Messverfahren zeigt eine erste Einschätzung, dass bei den möglichen neu formulierten Aufmerksamkeitsaspekten teilweise schon auf bestehende Testverfahren zurückgegriffen werden kann. Das gilt besonders für den Bereich „Reaktionsfähigkeit“ und den Bereich „Orientierung/räumliche Aufmerksamkeit“.

Für die möglichen neuen Bereiche „Aufmerksamkeitsteilung“, „Fokussierte Aufmerksamkeit“ und „Daueraufmerksamkeit“ (im Sinne von „Belastbarkeit“) müssen wahrscheinlich jedoch neue Test konstruiert werden.

Langfristig sollte, wie auf dem Expertenworkshop vorgeschlagen, vor einer faktischen Festlegung neuerer Aufmerksamkeitsbereiche in erster Linie

kriteriumsorientiert vorgegangen werden. Ausgehend von einer Anforderungsanalyse des Fahrverhaltens im Realverkehr sollte geprüft werden, welche der neueren wissenschaftlichen Aufmerksamkeitsaspekte in welcher Form mindestens vorliegen müssen, um die Fahraufgabe erfüllen zu können.

Die empirische Studie, die den zweiten Teil des Forschungsprojektes darstellt, ergab, dass die Messergebnisse der Reaktionszeiten eines in Deutschland zugelassenen Computertests von den verwendeten PC-Systemen, Monitoren und Softwarekonfigurationen abhängig sind. Die durchschnittliche Abweichung der gemessenen von den objektiven Reaktionszeiten beträgt 20,7 ms zu Ungunsten des Probanden. Das scheint für den Bereich „Reaktionszeit“ zunächst keine elementare Ungenauigkeit zu sein. Wie die weitere Analyse jedoch zeigt, kann diese Zeitdifferenz eine deutliche Verschiebung bei der Einordnung in die Prozentrangwerte zu Ungunsten des Probanden für den zu messenden Bereich „Reaktionsfähigkeit“ bedeuten und könnte letztlich dazu führen, dass die Mindestanforderung im Bereich der „Reaktionsfähigkeit“ als nicht erfüllt angesehen wird, obwohl objektiv entsprechende Leistungen vorliegen. Dies gilt insbesondere für Probanden, die eher als leistungsschwächer einzuordnen sind. Zur genaueren Beurteilung der im Rahmen einer Fahreignungsuntersuchung gemessenen Reaktionszeit (gerade in den etwas leistungsschwächeren Prozentrang-Bereichen) ist es daher unbedingt erforderlich, den maximal möglichen Messfehler eines computergestützten Messverfahrens anzugeben. Um eine gemessene Leistung darüber hinaus in den richtigen Kontext stellen zu können, muss es zusätzlich möglich sein, die vollständigen Normwerttabellen – für jeden einzelnen Prozentrang mit dem entsprechenden Konfidenzintervall dargestellt – im Überblick einsehen zu können. Nur so kann ein Untersucher entscheiden, ob eine Leistung eher am unteren oder eher am oberen Rand einer bestimmten Leistungskategorie einzuordnen ist.

12 Literatur

- AGILENT Technologis Inc. [Hrsg.]. (2002): User's Guide. Agilent 33120A 15 MHz Function/Arbitrary Waveform Generator. Publication Number 33120-90006, Edition 6, March 2002
- AKINWUNTAN, A. E., FEYS, H., DE, W. W., BATEN, G., ARNO, P. & KIEKENS, C. (2006): Prediction of driving after stroke: a prospective study. *Neurorehabilitation and Neural Repair*, 20, 417-423
- AKINWUNTAN, A. E., FEYS, H., DEWEERDT, W., PAUWELS, J., BATEN, G. & STRYPSTEIN, E. (2002): Determinants of driving after stroke. *Archives of Physical Medicine and Rehabilitation*, 83, 334-3410
- ANSTEY, K. J., WOOD, J., LORD, S. & WALKER, J. G. (2005): Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychological Review*, 25, 45-65
- BORTZ, J. & DÖRING, N. (2003): *Forschungsmethoden und Evaluation*. Berlin: Springer
- BROUWER, W. H. (2002): Attention and driving: A cognitive neuropsychological approach. In M. LECLERCQ & P. ZIMMERMANN: *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation*. Hove: Psychology Press
- BROWN, L. B., STERN, R. A., CAHN-WEINER, D. A., ROGERS, B., MESSER, M. A., LANNON, M. C., MAXWELL, C., SOUZA, T., WHITE, T. & OTT, B. R. (2005): Driving Scenes test of the Neuropsychological Assessment Battery (NAB) and on-road driving performance in aging and very mild dementia. *Archives of Clinical Neuropsychology*, 20, 209-215
- BUKASA, B., CHRIST, R., PONOCNY-SELIGER, E., SMUC, M. & WENNINGER, U. (2003a): Zweite Fahrverhaltensvalidierung der ART2020-Testverfahren
- BUKASA, B. & PIRINGER, A. (2001): Validierungsstudien zur Überprüfung der Aussagekraft von Leistungstests für die Fahreignungsbegutachtung. *Psychologie in Österreich*, 3, 187-194
- BUKASA, B., CHRIST, R., PONOCNY-SELIGER, E., SMUC, M. & WENNINGER, U. (2003b): Validitätsüberprüfung verkehrspsychologischer Leistungstests für die Fahreignungsbegutachtung

- tung. Zeitschrift für Verkehrssicherheit, 4, 191-197
- Bundesanstalt für Straßenwesen (2000): Begutachtungs-Leitlinien zur Kraftfahrereignung. Berichte der Bundesanstalt für Straßenwesen. Mensch und Sicherheit, Heft M 115. Bremerhaven: Wirtschaftsverlag NW
- CALLEJAS, A., LUPIANEZ, J. & TUDELA, P. (2004): The three attentional networks: on their independence and interactions. *Brain and Cognition*, 54, 225-227
- CORBETTA, M., KINCADE, J. M., OLLINGER, J. M., McAVOY, M. P. & SHULMAN, G. L. (2000): Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3, 292-297
- de RAEDT, R. & PONJAERT-KRISTOFFERSEN, I. (2000): The relationship between cognitive/neuropsychological factors and car driving performance in older adults. *Journal of the American Geriatric Society*, 48, 1664-1668
- DIN (2002). DIN 33430: Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen. Berlin: Beuth Verlag GmbH
- FAN, J., McCANDLISS, B. D., SOMMER, T., RAZ, A. & POSNER, M. I. (2002): Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14, 340-347
- FERNANDEZ-DUQUE, D. & POSNER, M. I. (2001): Brain imaging of attentional networks in normal and pathological states. *Journal of Clinical and Experimental Neuropsychology*, 23, 74-93
- FIMM, B., ZAHN, R., MULL, M., KEMENY, S., BUCHWALD, F., BLOCK, F. & SCHWARZ, M. (2001): Asymmetries of visual attention after circumscribed subcortical vascular lesions. *Journal of Neurology Neurosurgery and Psychiatry*, 71, 652-657
- GITELMAN, D. R., NOBRE, A. C., PARRISH, T. B., LABAR, K. S., KIM, Y. H., MEYER, J. R. & MESULAM, M. (1999): A large-scale distributed network for covert spatial attention: further anatomical delineation based on stringent behavioural and cognitive controls. *Brain*, 122 (Pt 6), 1093-1106
- GITELMAN, D. R., PARRISH, T. B., FRISTON, K. J. & MESULAM, M. M. (2002): Functional anatomy of visual search: regional segregations within the frontal eye fields and effective connectivity of the superior colliculus. *Neuroimage*, 15, 970-982
- GOLZ, D., HUCHLER, S., JÖRG, A. & KÜST, J. (2004): Beurteilung der Fahreignung. *Zeitschrift für Neuropsychologie*, 15, 157-167
- GRACE, J., AMICK, M. M., D'ABREU, A., FESTA, E. K., HEINDEL, W. C. & OTT, B. R. (2005): Neuropsychological deficits associated with driving performance in Parkinson's and Alzheimer's disease. *Journal of the International Neuropsychological Society*, 11, 766-775
- HANNEN, P., HARTJE, W. & SKRECZEK, W. (1998): Beurteilung der Fahreignung nach Hirnschädigung: Neuropsychologische Diagnostik und Fahrprobe. *Nervenarzt*, 69, 864-872
- HASHER, L., STOLTZFUS, E. R., ZACKS, R. T. & RYPMA, B. (1991): Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 163-169
- HEIN, G. & SCHUBERT, T. (2004): Aging and input processing in dual-task situations. *Psychology and Aging*, 19, 416-432
- HELD, T., LAMBERTI, G. & KUBITZKI, J. (1993): Psychose und Fahreignung – Entwicklung von Kriterien zur klinischen Beurteilung. *Die Rehabilitation*, 32, 155-161
- HILDEBRANDT, H., GIESSELMANN, H. & SACHSENHEIMER, W. (1999): Visual search and visual target detection in patients with infarctions of the left or right posterior or the right middle brain artery. *Journal of Clinical and Experimental Neuropsychology*, 21, 94-107
- HOMMEL, B., LI, K. Z. & LI, S. C. (2004): Visual search across the life span. *Developmental Psychology*, 40, 545-558
- KARNATH, H. O., FERBER, S. & HIMMELBACH, M. (2001): Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature*, 411, 950-953
- KARNER, T. & BIEHL, B. (2001): Über die Zusammenhänge verschiedener Versionen von Leistungstests im Rahmen der verkehrspsychologischen Diagnostik. *Zeitschrift für Verkehrssicherheit*, 2, 53-63

- KONDO, H., OSAKA, N. & OSAKA, M. (2004): Cooperation of the anterior cingulate cortex and dorsolateral prefrontal cortex for attention shifting. *Neuroimage*, 23, 670-679
- KROLL, G., KAISER, A., KRONE, M., MONNING, M., GRIESE, H., MACEK, C. & HARTJE, W. (2003): Die praktische Fahrprobe im mittleren und höheren Lebensalter. *Zeitschrift für Neuropsychologie*, 14, 81-87
- LIENERT, G. A. & RAATZ, U. (1998): Testaufbau und -analyse. Weinheim: Psychologie Verlags Union
- MARSHALL, S. C., MOLNAR, F., MAN-SON-HING, M., BLAIR, R., BROSSEAU, L., FINESTONE, H. M., LAMOTHE, C., KORNER-BITENSKY, N. & WILSON, K. G. (2007): Predictors of driving ability following stroke: a systematic review. *Topica in Stroke Rehabilitation*, 14, 98-114
- MILHAM, M. P., BANICH, M. T., CLAUS, E. D. & COHEN, N. J. (2003): Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *Neuroimage*, 18, 483-493
- MILHAM, M. P., ERICKSON, K. I., BANICH, M. T., KRAMER, A. F., WEBB, A., WSZALEK, T. & COHEN, N. J. (2002): Attentional control in the aging brain: insights from an fMRI study of the stroop task. *Brain and Cognition*, 49, 277-296
- NOBRE, A. (2004): Probing the Flexibility of Attentional Orienting in the Human Brain. In M. I. POSNER: *Cognitive Neuroscience of Attention*. New York: Guilford Press
- NOBRE, A. C. (2001): Orienting attention to instants in time. *Neuropsychologia*, 39, 1317-1328
- NOVACK, T. A., BANOS, J. H., ALDERSON, A. L., SCHNEIDER, J. J., WEED, W., BLANKENSHIP, J. & SALISBURY, D. (2006): UFOV performance and driving ability following traumatic brain injury. *Brain Injury*, 20, 455-461
- OLSON, E., STARK, M. & CHATTERJEE, A. (2003): Evidence for a unimodal somatosensory attention system. *Experimental Brain Research*, 151, 15-23
- POSNER, M. I. & RAICHLE, M. E. (1994): *Images of mind*. New York: WH Freeman
- REGER, M. A., WELSH, R. K., WATSON, G. S., CHOLERTON, B., BAKER, L. D. & CRAFT, S. (2004): The relationship between neuropsychological functioning and driving ability in dementia: a meta-analysis. *Neuropsychology*, 18, 85-93
- SHALLICE, T. (1988): *From Neuropsychology to mental structure*. Cambridge: University Press
- SINGH, R., PENTLAND, B., HUNTER, J. & PROVAN, F. (2007): Parkinson's disease and driving ability. *Journal of Neurology, Neurosurgery and Psychiatry*, 78, 363-366
- SMUC, M., CHRIST, R. & GATSCHA, M. (2006): SAF (System zur Analyse des Fahrverhaltens)
- SOMERS, D. C., DALE, A. M., SEIFFERT, A. E. & TOOTELL, R. B. (1999): Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of National Academy of Science U.S.A.*, 96, 1663-1668
- SOMMER, M., ARENDASY, M., SCHUHFRIED, G. & LITZENBERGER, M. (2005): Diagnostische Unterscheidbarkeit unfallfreier und mehrfach unfallbelasteter Kraftfahrer mit Hilfe nicht-linearer Auswertungsmethoden. *Zeitschrift für Verkehrssicherheit*, 2, 82-86
- SOMMER, M. & HÄUSLER, J. (2006): Kriteriumsvalidität des Expertensystems Verkehr. *Zeitschrift für Verkehrssicherheit*, 2, 83-89
- SPENCE, C. & DRIVER, J. (2004): *Crossmodal space & crossmodal attention*. Oxford: University Press
- STURM, W., DE, S. A., KRAUSE, B. J., SPECHT, K., HESSELMANN, V., RADERMACHER, I., HERZOG, H., TELLMANN, L., MULLER-GARTNER, H. W. & WILLMES, K. (1999): Functional anatomy of intrinsic alertness: evidence for a fronto-parietal-thalamic-brainstem network in the right hemisphere. *Neuropsychologia*, 37, 797-805
- STURM, W. & WILLMES, K. (2001): On the functional neuroanatomy of intrinsic and phasic alertness. *Neuroimage*, 14, S76-S84
- Testkuratorium der Föderation Deutscher Psychologenv Verbände (1986): Mitteilung. Beschreibung der einzelnen Kriterien für die Testbeurteilung. *Diagnostica*, 32, 358-360

- TSANG, P. S. & SHANER, T. L. (1998): Age, attention, expertise, and time-sharing performance. *Psychology and Aging*, 13, 323-347
- ULRICH, R. & GIRAY, M. (1989): Time resolution of clocks: Effects on reaction time measurement – Good News for bad clocks. *British Journal of Mathematical and Statistical Psychology*, 42, S. 1-12
- UTTL, B. & PILKENTON-TAYLOR, C. (2001): Letter cancellation performance across the adult life span. *The Clinical Neuropsychologist*, 15, 521-530
- Van ZOMEREN, A. H. & BROUWER, W. H. (1994): *Clinical Neuropsychology of Attention*. New York: Oxford University Press
- WARTENBURG, F. (ohne Jahresangabe): Does a clock resolution of 16 ms cause problems when measuring movement times around 100 ms? Available from wartenbe@informatik.uni-hamburg.de
- ZIMMERMANN, P. & FIMM, B. (2002a): A test battery for attentional performance. In M. LECLERCQ & P. ZIMMERMANN: *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation*. Hove: Psychology Press
- ZIMMERMANN, P. & FIMM, B. (2002b): Testbatterie zur Aufmerksamkeitsprüfung. Version 1.7. Herzogenrath: Psytest
- ZIMMERMANN, P. & LECLERCQ, M. (2002): Neuropsychological aspects of attentional functions and disturbances. In: M. LECLERCQ & P. ZIMMERMANN: *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation*. Hove: Psychology Press
- 12.1 Ergänzende Literatur „psychometrische Leistungstests zur Fahr-eignung und Fahrprobe“**
- ADLER, G., ROTTUNDA, S. & DYSKEN, M. (2005): The older driver with dementia: an updated literature review. *Journal of Safety Research*, 36, 399-407
- AHLGREN, E., LUNDQVIST, A., NORDLUND, A., AREN, C. & RUTBERG, H. (2003): Neurocognitive impairment and driving performance after coronary artery bypass surgery. *European Journal of Cardio-Thoracic Surgery*, 23, 334-340
- AK 1.1 „Fahr-, Steuer- und Überwachungstätigkeiten“ im Ausschuss Arbeitsmedizin des HVBG. Anwendungshinweise für den berufsgenossenschaftlichen Grundsatz „Fahr-, Steuer- und Überwachungstätigkeiten“ (G 25). [www.hvbg.de/d/bgz/praevaus/amed/bg_grund/g_25pdf.pdf; 13.02.08]
- AKINWUNTAN, A. E., FEYS, H., DEWEERDT, W., PAUWELS, J., BATEN, G. & STRYPSTEIN, E. (2002): Determinants of driving after stroke. *Archives of Physical Medicine and Rehabilitation*, 83, 334-341
- AKINWUNTAN, A. E., DEWEERDT, W., FEYS, H., BATEN, G., ARNO, P. & KIEKENS, C. (2003): Reliability of a road test after stroke. *Archives of Physical Medicine and Rehabilitation*, 84, 1792-1796
- AKINWUNTAN, A. E., De, W. W., FEYS, H., BATEN, G., ARNO, P. & KIEKENS, C. (2005): The validity of a road test after stroke. *Archives of Physical Medicine and Rehabilitation*, 86, 421-426
- AKINWUNTAN, A. E., De, W. W., FEYS, H., PAUWEIS, J., BATEN, G., ARNO, P. & KIEKENS, C. (2005): Effect of simulator training on driving after stroke: a randomized controlled trial. *Neurology*, 65, 843-850
- AKINWUNTAN, A. E., FEYS, H., De, W. W., BATEN, G., ARNO, P. & KIEKENS, C. (2006): Prediction of driving after stroke: a prospective study. *Neurorehabilitation and Neural Repair*, 20, 417-423
- ARMELANG, M. & SCHMIDT-ATZERT, L. (2006): *Psychologische Diagnostik und Intervention*. Berlin: Springer
- ANSTEY, K. J., WOOD, J., LORD, S. & WALKER, J. G. (2005): Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review*, 25, 45-65
- ARNEDT, J. T., GEDDES, M. A. & McLEAN, 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

- BALDOCK, M. R., MATHIAS, J. L., McLEAN, A. J. & BERNDT, A. (2006): Self-regulation of driving and its relationship to driving ability among older adults. *Accident Analysis and Prevention*, 38, 1038-1045
- BALL, K., OWSLEY, C., SLOANE, M. E., ROENKER, D. L. & BRUNI, J. R. (1993): Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology and Visual Science*, 34, 3110-3123
- BALL, K., OWSLEY, C., STALVEY, B., ROENKER, D. L., SLOANE, M. E. & GRAVES, M. (1998): Driving avoidance and functional impairment in older drivers. *Accident Analysis and Prevention*, 30, 313-322
- BALL, K. K., ROENKER, D. L., WADLEY, V. G., EDWARDS, J. D., ROTH, D. L., MCGWIN, G. Jr., RALEIGH, R., JOYCE, J. J., CISELL, G. M. & DUBE, T. (2006): Can high-risk older drivers be identified through performance-based measures in a Department of Motor Vehicles setting? *Journal of the American Geriatric Society*, 54, 77-84
- BARCS, G., VITRAI, J. & HALASZ, P. (1997): Investigation of vehicle driving ability in two diagnostic groups of epileptic patients with special neuropsychological approach. *Medicine and Law*, 16, 277-287
- BERG, M. & SCHUBERT, W. (1999): Das thematische Testsystem „Corporal“ zur Erfassung von Funktionen der Aufmerksamkeit. *Zeitschrift für Verkehrssicherheit*, 2, 74-81
- BERGHAUS, G. (2002): Methadon-Substitution und Fahrsicherheit/Fahreignung – Überblick ueber die experimentellen Forschungsergebnisse. *Suchtmedizin in Forschung und Praxis*, 4, 169-174
- BERKOVIC, S. F. (2001): Epilepsy syndromes: effects on cognition, performance and driving ability. *Medicine and Law*, 20, 547-551
- BIELIAUSKASS, L. A. (2005): Neuropsychological assessment of geriatric driving competence. *Brain Injury*, 19, 221-226
- BORTZ, J. & DÖRING, N. (2003): *Forschungsmethoden und Evaluation*. Berlin: Springer
- BOWERS, A., PELI, E., ELGIN, J., MCGWIN, G. Jr. & OWSLEY, C. (2005): On-road driving with moderate visual field loss. *Optometry and Vision Science*, 82, 657-667
- BRADSTAETTER, C. (1992): The validity of traffic psychology testing procedures in assessing the driving aptitude of the elderly. *Psychologie in-Oesterreich*, 122, 25-30
- BRENNER-HARTMANN, J. & BUKASA, B. (2001): Psychologische Leistungsüberprüfung bei der Fahreignungsbegutachtung. *Zeitschrift für Verkehrssicherheit*, 1, 1-8
- BROOKS, N. & HAWLEY, C. A. (2005): Return to driving after traumatic brain injury: a British perspective. *Brain Injury*, 19, 165-175
- BROUWER, W. H. (2002): Attention and driving: A cognitive neuropsychological approach. In: M. LECLERCQ & P. ZIMMERMANN (eds.). *Applied Neuropsychology of Attention: Theory, Diagnosis and Rehabilitation*. Hove: Psychology Press
- BROWN, L. B. & OTT, B. R. (2004): Driving and dementia: a review of the literature. *Journal of the Geriatric Psychiatry and Neurology*, 17, 232-240
- BROWN, L. B., OTT, B. R., PAPANDONATOS, G. D., SUI, Y., READY, R. E. & MORRIS, J. C. (2005): Prediction of on-road driving performance in patients with early Alzheimer's disease. *Journal of the American Geriatric Society*, 53, 94-98
- BROWN, L. B., STERN, R. A., CAHN-WEINER, D. A., ROGERS, B., MESSER, M. A., LANNON, M. C., MAXWELL, C., SOUZA, T., WHITE, T. & OTT, B. R. (2005): Driving Scenes test of the Neuropsychological Assessment Battery (NAB) and on-road driving performance in aging and very mild dementia. *Archives of Clinical Neuropsychology*, 20, 209-215
- BRUNNAUER, A., LAUX, G., GEIGER, E., SOYKA, M. & MOLLER, H. J. (2006): Antidepressants and driving ability: results from a clinical study. *Journal of Clinical Psychiatry*, 67, 1776-1781
- BRUNNAUER, A., LAUX, G. & GEIGER, E. (2007): Driving Ability of Depressive Patients under Mirtazapine and Reboxetine. *Psychiatrische Praxis*, 34, 84-86

- BUKASA, B. & PIRINGER, A. (2001): Validierungsstudien zur Überprüfung der Aussagekraft von Leistungstests für die Fahreignungsbegutachtung. *Psychologie in Österreich*, 3, 187-194
- BUKASA, B., CHRIST, R., PONOCNY-SELIGER, E., SMUC, M. & WENNINGER, U. (2003): Zweite Fahrverhaltensvalidierung der ART2020-Testverfahren
- BUKASA, B., CHRIST, R., PONOCNY-SELIGER, E., SMUC, M. & WENNINGER, U. (2003): Validitätsüberprüfung verkehrspsychologischer Leistungstests für die Fahreignungsbegutachtung. *Zeitschrift für Verkehrssicherheit*, 4, 191-197
- Bundesanstalt für Straßenwesen (2002): Begutachtungs-Leitlinien zur Kraftfahrereignung. *Berichte der Bundesanstalt für Straßenwesen. Mensch und Sicherheit*, Heft M 115. Bremerhaven: Wirtschaftsverlag NW
- BURGARD, E., SICK, C.-D., KISS, M. & HIPPEL, B. (2004): Fahreignung nach Hirnschädigung – Erfahrungen aus der klinischen Praxis der Begutachtung und Beratung. *Zeitschrift für Verkehrssicherheit*, 3, 145-150
- BURGARD, E. (2005): Fahrkompetenz im Alter: Die Aussagekraft diagnostischer Instrumente bei Senioren und neurologischen Patienten. Ludwig-Maximilians-Universität München
- CALLEJAS, A., LUPIANEZ, J. & TUDELA, P. (2004): The three attentional networks: on their independence and interactions. *Brain and Cognition*, 54, 225-227
- CHALOUPKA-RISSER, C. & RISSER, R. (2006): Verkehrstauglichkeit aus neuropsychologischer und verkehrspsychologischer Sicht. In J. LEHRNER, G. PUSSWALD, E. FERL, I. KRYSPIEXNER & W. STRUBREITHER (Hrsg). *Klinische Neuropsychologie*. Wien: Springer
- CHALOUPKA, C. & RISSER, R. (1995): Don't wait for accidents- Possibilities to assess risk in traffic by applying the 'Wiener Fahrprobe'. *Safety Science*, 19, 137-147
- CLAY, O. J., WADLEY, V. G., EDWARDS, J. D., ROTH, D. L., ROENKER, D. L. & BALL, K. K. (2005): Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: current and future implications. *Optometry and Vision Science*, 82, 724-731
- CORBETTA, M., KINCADE, J. M., OLLINGER, J. M., McAVOY, M. P. & SHULMAN, G. L. (2000): Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3, 292-297
- COX, D. J., PENBERTHY, J. K., MERKEL, R. L. & KOVATCHEV, B. (2002): Driving performance among adolescents with attention-deficit/hyperactivity disorder: Medication effects. *European Neuropsychopharmacology*, 12, 415-416
- CRABB, D. P., FITZKE, F. W., HITCHINGS, R. A. & VISWANATHAN, A. C. (2004): A practical approach to measuring the visual field component of fitness to drive. *British Journal of Ophthalmology*, 88, 1191-1196
- CUSHMAN, L. A. & COGLIANDRO, F. C. (1999): On-road driving post-stroke: Cognitive and other factors. *Archives of Clinical Neuropsychology*, 14, 799-795
- De RAEDT, R. & PONJAERT-KRISTOFFERSEN, I. (2000): The relationship between cognitive/neuropsychological factors and car driving performance in older adults. *Journal of the American Geriatric Society*, 48, 1664-1668
- Deutsche Gesellschaft für Psychologie (2006): *Aktuelle Mitteilungen*, Jahrgang 2, Ausgabe 4
- DIN (2002): DIN 33430. Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen. Berlin: Beuth Verlag GmbH
- DUCHEK, J. M., HUNT, L., BALL, K., BUCKLES, V. & MORRIS, J. C. (1997): The role of selective attention in driving and dementia of the Alzheimer type. *Alzheimer Disease and Associated Disorder*, 11 Suppl 1, 48-56
- DUCHEK, J. M., HUNT, L., BALL, K., BUCKLES, V. & MORRIS, J. C. (1998): Attention and driving performance in Alzheimer's disease. *Journal of Gerontology Series B. Psychological Sciences and Social Sciences*, 53, 130-141
- EDWARDS, J. D., VANCE, D. E., WADLEY, V. G., CISELL, G. M., ROENKER, D. L. & BALL, K. K. (2005): Reliability and validity of useful field of

- view test scores as administered by personal computer. *Journal of Clinical and Experimental Neuropsychology*, 27, 529-543
- EDWARDS, J. D., ROSS, L. A., WADLEY, V. G., CLAY, O. J., CROWE, M., ROENKER, D. L. & BALL, K. K. (2006): The useful field of view test: normative data for older adults. *Archives of Clinical Neuropsychology*, 21, 275-286
- ELKIN-FRANKSTON, S., LEBOWITZ, B. K., KAPUST, L. R., HOLLIS, A. M. & O'CONNOR, M. G. (2007): The use of the Color Trails Test in the assessment of driver competence: Preliminary report of a culture-fair instrument. *Archives of Clinical Neuropsychology*, 22, 613-635
- Fahrerlaubnis-Verordnung (FeV) (2006): Verordnung über die Zulassung von Personen zum Straßenverkehr
- FALKENSTEIN, M., HIELSCHER, H., DZIOBEK, I., SCHWARZAU, P., HOORMANN, J., SUNDERMAN, B. & HOHNSBEIN, J. (2001): Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport*, 12, 157-161
- FALKENSTEIN, M., HOORMANN, J. & HOHNSBEIN, J. (2001): Changes of error-related ERPs with age. *Experimental Brain Research*, 138, 258-262
- FAN, J., McCANDLISS, B. D., SOMMER, T., RAZ, A. & POSNER, M. I. (2002): Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14, 340-347
- FERMINA, S. & VEZ, M. (2005): Prevention of traffic accidents: The assessment of perceptual-motor alterations before obtaining a driving license. A longitudinal study of the first years of driving. *Brain Injury*, 19, 189-196
- FERNANDEZ-DUQUE, D. & POSNER, M. I. (2001): Brain imaging of attentional networks in normal and pathological states. *Journal of Clinical and Experimental Neuropsychology*, 23, 74-93
- FIMM, B. (1997): Mikroanalyse von Aufmerksamkeitsprozessen. In: S. GAUGGEL & G. KERKHOF (Hrsg.). *Fallbuch Neuropsychologie*. Göttingen: Hogrefe
- FIMM, B. (1998): Testrezension: Der Aufmerksamkeits-Belastungs-Test d2. *Report Psychologie*, 23, 147-153
- FIMM, B. (2002): Rezension: HEUBROCK, D. & PETERMANN, F. (2002): Aufmerksamkeitsdiagnostik. *Zeitschrift für Neuropsychologie*, 13, 255-256
- FIMM, B., ZAHN, R., MULL, M., KEMENY, S., BUCHWALD, F., BLOCK, F. & SCHWARZ, M. (2001): Asymmetries of visual attention after circumscribed subcortical vascular lesions. *Journal of Neurology, Neurosurgery and Psychiatry*, 71, 652-657
- FISCHER, M., BARKLEY, R. A., SMALLISH, L. & FLETCHER, K. (2007): Hyperactive children as young adults: driving abilities, safe driving behavior, and adverse driving outcomes. *Accident Analysis and Prevention*, 39, 94-105
- FISK, G. D., NOVACK, T., MENNEMEIER, M. & ROENKER, D. (2002): Useful field of view after traumatic brain injury. *Journal of the Head Trauma Rehabilitation*, 17, 16-25
- FISK, G. D. & MENNEMEIER, M. (2006): Common neuropsychological deficits associated with stroke survivors' impaired performance on a useful field of view test. *Perceptual and Motor Skills*, 102, 387-394
- FORMISANO, R., BIVONA, U., BRUNELLI, S., GIUSTINI, M., LONGO, E. & TAGGI, F. (2005): A preliminary investigation of road traffic accident rate after severe brain injury. *Brain Injury*, 19, 159-163
- FRIES, W. & NETZ, J. (2005): Leitlinie zur Beurteilung der Fahreignung bei neurologischen Erkrankungen = Guideline for the Assessment of Driving ability of patients with Neurological diseases. *Akt Neurol [sic!]*, 32, 342-350
- GALSKI, T., WILLIAMS, J. B. & EHLE, H. T. (2000): Effects of Opioids on Driving Ability. *Journal of Pain and Symptom Management*, 19, 200-208
- GEIGER, E., BRUNNAUER, A. & LAUX, G. (2007): Driving ability in schizophrenic patients – stability and course. *Psychiatrische Praxis*, 34, 87-89
- GEORGE, S., CLARK, M. & CROTTY, M. (2007): Development of the Adelaide driving self-efficacy scale. *Clinical Rehabilitation*, 21, 56-61

- GITELMAN, D. R., NOBRE, A. C., PARRISH, T. B., LABAR, K. S., KIM, Y. H., MEYER, J. R. & MESULAM, M. (1999): A large-scale distributed network for covert spatial attention: further anatomical delineation based on stringent behavioural and cognitive controls. *Brain*, 122 (Pt 6), 1093-1106
- GITELMAN, D. R., PARRISH, T. B., FRISTON, K. J. & MESULAM, M. M. (2002): Functional anatomy of visual search: regional segregations within the frontal eye fields and effective connectivity of the superior colliculus. *Neuroimage*, 15, 970-982
- GOLZ, D., HUCHLER, S., JÖRG, A. & KÜST, J. (2004): Beurteilung der Fahreignung. *Zeitschrift für Neuropsychologie*, 15, 157-167
- GRACE, J., AMICK, M. M., D'ABREU, A., FESTA, E. K., HEINDEL, W. C. & OTT, B. R. (2005): Neuropsychological deficits associated with driving performance in Parkinson's and Alzheimer's disease. *Journal of the International Neuropsychological Society*, 11, 766-775
- GSTALTER, H. (2005): Anforderungen der Fahraufgabe und tatsächliches Fahrverhalten. Ergebnisse eines Soll-/Ist-Vergleichs mit der neuen Analyseverfahren SAFE; Requirements of the driving task and drivers' actual behavior – Results of a comparison using SAFE, the new method of driver task analysis. *Zeitschrift für Verkehrssicherheit*, 51, 76-82
- HAMANN, K. F. (2002): Driving ability with vestibular lesions. *HNO*, 50, 1086-1088
- HANNEN, P. (1997): Fahrverhalten nach erworbenen Hirnschäden – eine empirische Studie. Universität Bielefeld
- HANNEN, P., HARTJE, W. & SKRECZEK, W. (1998): Beurteilung der Fahreignung nach Hirnschädigung: Neuropsychologische Diagnostik und Fahrprobe. *Nervenarzt*, 69, 864-872
- HANNEN, P. (1991): Beurteilung der Fahreignung hirngeschädigter Patienten; Diagnosis of driving ability in patients with brain damage. In: A COHEN & R. HIRSIG. *Fortschritte der Verkehrspsychologie 1990*. 30. BDP-Kongress für Verkehrspsychologie und Fortbildungsveranstaltung Rorschach (S. 316-322)
- HASHER, L., STOLTZFUS, E. R., ZACKS, R. T. & RYPMA, B. (1991): Age and inhibition. *Journal of Experimental Psychology, Learning, Memory and Cognition*, 17, 163-169
- HAYMES, S. A., LEBLANC, R. P., NICOLELA, M. T., CHIASSON, L. A. & CHAUHAN, B. C. (2007): Risk of falls and motor vehicle collisions in glaucoma. *Investigative Ophthalmology and Visual Science*, 48, 1149-1155
- HEBER, I. A., VALVODA, J. T., KUHLEN, T., STURM, W. & FIMM, B. (2004): Visuelle Aufmerksamkeitsausrichtung im virtuellen Raum. *Zeitschrift für Neuropsychologie*, 15, 121-121
- HEBER, I. A., VALVODA, J. T., KUHLEN, T., STURM, W. & FIMM, B. (2005): Asymmetries of visuo-spatial attention in virtual space after sleep deprivation. *Journal of Cognitive Neuroscience*, Suppl. 51
- HEIKKILA, V. M., TURKKA, J., KORPELAINEN, J., KALLANRANTA, T. & SUMMALA, H. (1998): Decreased driving ability in people with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, 64, 325-330
- HEIKKILA, V. M., KORPELAINEN, J., TURKKA, J., KALLANRANTA, T. & SUMMALA, H. (1999): Clinical evaluation of the driving ability in stroke patients. *Acta Neurologica Scandinavica*, 99, 349-355
- HEIKKILA, V. M. & KALLANRANTA, T. (2005): Evaluation of the driving ability in disabled persons: a practitioners' view. *Disability and Rehabilitation*, 27, 1029-1036
- HEIN, G. & SCHUBERT, T. (2004): Aging and input processing in dual-task situations. *Psychology and Aging*, 19, 416-432
- HELD, T., LAMBERTI, G. & KUBITZKI, J. (1993): Psychose und Fahreignung – Entwicklung von Kriterien zur klinischen Beurteilung. *Die Rehabilitation*, 32, 155-161
- HILDEBRANDT, H., GIESSELMANN, H. & SACHSENHEIMER, W. (1999): Visual search and visual target detection in patients with infarctions of the left or right posterior or the right middle brain artery. *Journal of Clinical and Experimental Neuropsychology*, 21, 94-107
- HOFFMAN, L., McDOWD, J. M., ATCHLEY, P. & DUBINSKY, R. (2005): The role of visual attention in predicting driving impairment in older adults. *Psychology and Aging*, 20, 610-622

- HOMMEL, B., LI, K. Z. & LI, S. C. (2004): Visual search across the life span. *Developmental Psychology*, 40, 545-558
- HORIKAWA, E., OKAMURA, N., TASHIRO, M., SAKURADA, Y., MARUYAMA, M., ARAI, H., YAMAGUCHI, K., SASAKI, H., YANAI, K. & ITOH, M. (2005): The neural correlates of driving performance identified using positron emission tomography. *Brain and Cognition*, 58, 166-171
- HUNT, L. A., MURPHY, C. F., CARR, D., DUCHEK, J. M., BUCKLES, V. & MORRIS, J. C. (1997): Environmental cueing may effect performance on a road test for drivers with dementia of the Alzheimer type. *Alzheimer Disease and Associated Disorder*, 11 Suppl. 1, 13-16
- HUNT, L. A., MURPHY, C. F., CARR, D., DUCHEK, J. M., BUCKLES, V. & MORRIS, J. C. (1997): Reliability of the Washington University Road Test. A performance-based assessment for drivers with dementia of the Alzheimer type. *Archives of Neurology*, 54, 707-712
- INNES, H., JONES, R., ANDERSON, T., RYMPLE-ALFORD, J., HAYES, S., HOLLOBON, S., SEVERINSEN, J., SMITH, G. & NICHOLLS, A. (2005): Prediction of Driving Ability in Persons with Brain Disorders using Sensory-Motor and Cognitive Tests. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society Conference*, 5, 5439-5442
- JACOBSHAGEN, W. (2005): Leistungstests für Berufskraftfahrer. *Zeitschrift für Verkehrssicherheit*, 3, 125-131
- JANKE, M. K. & EBERHARD, J. W. (1998): Assessing medically impaired older drivers in a licensing agency setting. *Accident Analysis and Prevention*, 30, 347-361
- JANKE, M. K. (2001): Assessing older drivers: Two studies. *Journal of Safety Research*, 32, 43-74
- KAGERER, S., WINTER, C., MOLLER, H. J. & SOYKA, M. (2003): Effects of haloperidol and atypical neuroleptics on psychomotor performance and driving ability in schizophrenic patients. Results from an experimental study. *Neuropsychobiology*, 47, 212-218
- KARNATH, H. O., FERBER, S. & HIMMELBACH, M. (2001): Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature*, 411, 950-953
- KARNER, T. & BIEHL, B. (2001): Über die Zusammenhänge verschiedener Versionen von Leistungstests im Rahmen der verkehrspsychologischen Diagnostik. *Zeitschrift für Verkehrssicherheit*, 2, 53-63
- KARNER, T. & NEUWIRTH, W. (2001): Die Bedeutung der peripheren Wahrnehmung in der verkehrspsychologischen Untersuchung. *Psychologie in Österreich*, 21, 183-186
- KAUFMANN, R. M., FREY, R., BATTISTA, H. J. & KASPER, S. (2004): Flunitrazepam and driving ability. *Fortschritte der Neurologie-Psychiatrie*, 72, 503-515
- KELLER, M. (2007): Weiterfahren oder nicht? Verkehrspsychologie in der neurologischen Rehabilitation. *Psychoscope*, 25, 10-11
- KERSTING, M. (2004): Anforderungen an die Qualität der Verfahren. In: L. F. HORNKE & U. WINTERFELD (Hrsg.). *Eignungsbeurteilungen auf dem Prüfstand: DIN 33430 zur Qualitätssicherung*. Heidelberg: Spektrum Akademischer
- KLEBELSBERG, D. (1995): Optimierungsprobleme bei der Beurteilung der Fahreignung. *Zeitschrift für Verkehrssicherheit*, 1, 30-33
- KLEMENJAK, W. (2006): Akute und chronische Beeinträchtigungen von Kraftfahrern. *Zeitschrift für Verkehrssicherheit*, 2, 62-63
- KONDO, H., OSAKA, N. & OSAKA, M. (2004): Cooperation of the anterior cingulate cortex and dorsolateral prefrontal cortex for attention shifting. *Neuroimage*, 23, 670-679
- KORTELING, J. E. & KAPTEIN, N. A. (1996): Neuropsychological driving fitness tests for brain-damaged subjects. *Archives of Physical Medicine and Rehabilitation*, 77, 138-146
- KOTTERBA, S., ORTH, M., EREN, E., FANGERAU, T. & SINDERN, E. (2003): Assessment of driving performance in patients with relapsing-remitting multiple sclerosis by a driving simulator. *European Neurology*, 50, 160-164
- KOTTERBA, S., MUELLER, N., LEIDAG, M., WIDDIG, W., RASCHE, K., MALIN, J. P., SCHULTZE-WERNINGHAUS, G. & ORTH, M. (2004): Comparison of driving simulator performance and neuropsychological testing in

- narcolepsy. *Clinical Neurology and Neurosurgery*, 106, 275-279
- KOTTERBA, S., WIDDIG, W., BRYLAK, S. & ORTH, M. (2005): Driving after cerebral ischemia – a driving simulator investigation. *Wiener Medizinische Wochenschrift.*, 155, 348-353
- KRAMM, T. (2000): Normierungsstudie einer Aufmerksamkeits- Testbatterie (Vigilanz, Signal-Detection und Cognitrone des Wiener Testsystems der Dr. G. Schuhfried Ges. m. b. H.). Ruhr-Universität Bochum
- KRESS, H. G. & KRAFT, B. (2005): Opioid medication and driving ability. *European Journal of Pain*, 9, 141-144
- KROLL, G., KAISER, A., KRONE, M., MONNING, M., GRIESE, H., MACEK, C. & HARTJE, W. (2003): Die praktische Fahrprobe im mittleren und höheren Lebensalter. *Zeitschrift für Neuropsychologie*, 14, 81-87
- KUBINGER, K. D. & JÄGER, R. S. (Hrsg.) (2003): Schlüsselbegriffe der Psychologischen Diagnostik. Weinheim: Beltz/PVU
- KUBINGER, K. D. (2006): Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens. Göttingen: Hogrefe
- KUBITZKI, J. (1992): Psychose und Fahreignung; Entwicklung von Kriterien zur klinischen Beurteilung; Psychosis and driving ability. Development of criteria for clinical evaluation. *Forschungsberichte*, Nr. 257)
- Kuratorium ZNS (2007): Neurologie und Verkehrssicherheit
- KURZTHALER, I., HUMMER, M., MILLER, C., SPERNER-UNTERWEGER, B., GUNTHER, V., WECHDORN, H., BATTISTA, H. J. & FLEISCH-HACKER, W. W. (1999): Effect of cannabis use on cognitive functions and driving ability. *Journal of Clinical Psychiatry*, 60, 395-399
- LAUX, G. (2002): Psychische Störungen und Fahrtauglichkeit – Eine Übersicht. *Der Nervenarzt*, 73, 231-238
- LEE, H. C., CAMERON, D. & LEE, A. H. (2003): Assessing the driving performance of older adult drivers: on-road versus simulated driving. *Accident Analysis and Prevention*, 35, 797-803
- LEON-CARRION, J., DOMINGUEZ-MORALES, M. R., BARROSO, J. M. & MARTIN, Y. (2005): Driving with cognitive deficits: neuro rehabilitation and legal measures are needed for driving again after severe traumatic brain injury. *Brain Injury*, 19, 213-219
- LEW, H. L., POOLE, J. H., LEE, E. H., JAFFE, D. L., HUANG, H. C. & BRODD, E. (2005): Predictive validity of driving-simulator assessments following traumatic brain injury: a preliminary study. *Brain Injury*, 19, 177-188
- LIENERT, G. A. & RAATZ, U. (1998): Testaufbau und -analyse. Weinheim: Psychologie Verlags Union
- MAAG, F. (1992): Fahrtauglichkeit; Theorie und Praxis der Fahrprobe: Verkehrsmedizinische Erfahrungen und ihr Stellenwert; Driving aptitude. Theory and practice of the driving sample. Experiences from traffic medicine. Bern: Huber
- MAAS, R., VENTURA, R., KRETZSCHMAR, C., AYDIN, A. & SCHUCHERT, A. (2003): Syncope, driving recommendations, and clinical reality: survey of patients. *BMJ (British Medical Journal)*, 326, 21
- MacMILLAN, R. (2005): Testing seniors' driving ability. *Canadian Family Physician*, 51, 655
- MAES, V., GRENEZ, O., CHARLIER, C., SMET, H., VERSTRAETE, A. & WENNIG, R. (1999): Classification of medicines according to their influence on driving ability. *Acta Clinica Belgica Supplementum* 1, 82-88
- MARCOTTE, T. D., WOLFSON, T., ROSENTHAL, T. J., HEATON, R. K., GONZALEZ, R., ELLIS, R. J. & GRANT, I. (2004): A multimodal assessment of driving performance in HIV infection. *Neurology*, 63, 1417-1422
- MARCOTTE, T. D., LAZZARETTO, D., SCOTT, J. C., ROBERTS, E., WOODS, S. P. & LETENDRE, S. (2006): Visual attention deficits are associated with driving accidents in cognitively-impaired HIV-infected individuals. *Journal of Clinical and Experimental Neuropsychology*, 28, 13-28
- MAROTTOLI, R. A. & RICHARDSON, E. D. (1998): Confidence in, and self-rating of, driving ability among older drivers. *Accident, analysis and Prevention*, 30, 331-336

- MAROTTOLI, R. A., ALLORE, H., ARAUJO, K. L., IANNONE, L. P., ACAMPORA, D., GOTTSCHALK, M., CHARPENTIER, P., KASL, S. & PEDUZZI, P. (2007): A randomized trial of a physical conditioning program to enhance the driving performance of older persons. *Journal of General International Medicine*, 22, 590-597
- MARSHALL, S. C., MOLNAR, F., MAN-SON-HING, M., BLAIR, R., BROSSEAU, L., FINESTONE, H. M., LAMOTHE, C., KORNER-BITENSKY, N. & WILSON, K. G. (2007): Predictors of driving ability following stroke: a systematic review. *Topics in Stroke Rehabilitation*, 14, 98-114
- MAZER, B. L., SOFER, S., KORNER-BITENSKY, N. & GELINAS, I. (2001): Use of the UFOV to evaluate and retrain visual attention skills in clients with stroke: a pilot study. *The American Journal of Occupational Therapy*, 55, 552-557
- MAZER, B. L., SOFER, S., KORNER-BITENSKY, N., GELINAS, I., HANLEY, J. & WOOD-DAUPHINEE, S. (2003): Effectiveness of a visual attention retraining program on the driving performance of clients with stroke. *Archives of Physical Medicine and Rehabilitation*, 84, 541-550
- McGWIN, G., Jr., OWSLEY, C. & BALL, K. (1998): Identifying crash involvement among older drivers: agreement between self-report and state records. *Accident, Analysis and Prevention*, 30, 781-791
- McKENNA, P., JEFFERIES, L., DOBSON, A. & FRUDE, N. (2004): The use of a cognitive battery to predict who will fail an on-road driving test. *The British Journal of Clinical Psychology*, 43, 325-336
- MEINDORFNER, C., KORNER, Y., MOLLER, J. C., STIASNY-KOLSTER, K., OERTEL, W. H. & KRUGER, H. P. (2005): Driving in Parkinson's disease: mobility, accidents, and sudden onset of sleep at the wheel. *Movement Disorder*, 20, 832-842
- MILHAM, M. P., ERICKSON, K. I., BANICH, M. T., KRAMER, A. F., WEBB, A., WSZALEK, T. & COHEN, N. J. (2002): Attentional control in the aging brain: insights from an fMRI study of the stroop task. *Brain and Cognition*, 49, 277-296
- MILHAM, M. P., BANICH, M. T., CLAUS, E. D. & COHEN, N. J. (2003): Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *Neuroimage*, 18, 483-493
- MOLLER, J. C., STIASNY, K., HARGUTT, V., CASSEL, W., TIETZE, H., PETER, J. H., KRUGER, H. P. & OERTEL, W. H. (2002): Evaluation of sleep and driving performance in six patients with Parkinson's disease reporting sudden onset of sleep under dopaminergic medication: a pilot study. *Movement Disorder*, 17, 474-481
- MORIKE, K. & GLEITER, C. (2003): Medicinal drugs and automobile driving ability. *Therapeutische Umschau*, 60, 347-354
- MÖRIKE, K. & GLEITER, C. H. (2002): Beeinflussung der Fahrtüchtigkeit durch Arzneimittel. *Der Internist*, 43, 889-898
- MYERS, R. S., BALL, K. K., KALINA, T. D., ROTH, D. L. & GOODE, K. T. (2000): Relation of useful field of view and other screening tests to on-road driving performance. *Perceptual Motor Skills*, 91, 279-290
- NOBRE, A. (2004): Probing the Flexibility of Attentional Orienting in the Human Brain. In: M. I. POSNER (Hrsg.). *Cognitive Neuroscience of Attention*. New York: Guilford Press
- NOBRE, A. C. (2001): Orienting attention to instants in time. *Neuropsychologia*, 39, 1317-1328
- NOVACK, T. A., BANOS, J. H., ALDERSON, A. L., SCHNEIDER, J. J., WEED, W., BLANKENSHIP, J. & SALISBURY, D. (2006): UFOV performance and driving ability following traumatic brain injury. *Brain Injury*, 20, 455-461
- O'JILE, J., GOUVIER, W., TODD, M., SANTULLI, R., SANTA MARIA, M., PINKSTON, J. & JOHNSON, S. (1999): Predicting driving ability in Alzheimer's disease. *Archives of Clinical Neuropsychology*, 14, 59-795
- OLSON, E., STARK, M. & CHATTERJEE, A. (2003): Evidence for a unimodal somatosensory attention system. *Experimental Brain Research*, 151, 15-23
- ORTH, M., DUCHNA, H. W., LEIDAG, M., WIDDIG, W., RASCHE, K., BAUER, T. T., WALTHER, J. W., De, Z. J., MALIN, J. P., SCHULTZE-WERNINGHAUS, G. & KOTTERBA, S. (2005): Driving simulator and neuropsychological [cor-

- rected] testing in OSAS before and under CPAP therapy. *European Respiratory Journal*, 26, 898-903
- OSTENDORF, G. M. (2006): Automobile driving simulation for diagnosis and therapy of driving fitness in neurologic disorders. *Versicherungsmedizin*, 58, 36-37
- OSWANSKI, M. F., SHARMA, O. P., RAJ, S. S., VASSAR, L. A., WOODS, K. L., SARGENT, W. M. & PITOCK, R. J. (2007): Evaluation of two assessment tools in predicting driving ability of senior drivers. *American Journal of Physical Medicine and Rehabilitation*, 86, 190-199
- OTT, B. R., HEINDEL, W. C., WHELIHAN, W. M., CARON, M. D., PIATT, A. L. & DICARLO, M. A. (2003): Maze test performance and reported driving ability in early dementia. *Journal of Geriatric Psychiatry and Neurology*, 16, 151-155
- OTT, B. R., ANTHONY, D., PAPANDONATOS, G. D., D'ABREU, A., BUROCK, J., CURTIN, A., WU, C. K. & MORRIS, J. C. (2005): Clinician assessment of the driving competence of patients with dementia. *Journal of American Geriatric Society*, 53, 829-833
- OWSLEY, C., MCGWIN, G., Jr. & BALL, K. (1998): Vision impairment, eye disease, and injurious motor vehicle crashes in the elderly. *Ophthalmic Epidemiology*, 5, 101-113
- OWSLEY, C., BALL, K., MCGWIN, G., Jr., SLOANE, M. E., ROENKER, D. L., WHITE, M. F. & OVERLEY, E. T. (1998): Visual processing impairment and risk of motor vehicle crash among older adults. *Journal of the American Medical Association*, 279, 1083-1088
- PARASURAMAN, R. & NESTOR, P. (1993): Attention and driving. Assessment in elderly individuals with dementia. *Clinics in Geriatric Medicine*, 9, 377-387
- PIETRAPIANA, P., TAMIETTO, M., TORRINI, G., MEZZANATO, T., RAGO, R. & PERINO, C. (2005): Role of premorbid factors in predicting safe return to driving after severe TBI. *Brain Injury*, 19, 197-211
- POSCHADEL, S., ROENSCH-HASSELHORN, B. & SOMMER, S.-M. (2006): Eignungsbegutachtung zur Mobilitätsförderung älterer Kraftfahrer – Ergebnisse des EU-Projekts AGILE. *Zeitschrift für Verkehrssicherheit*, 52, 13-18
- POSNER, M. I. & RAICHLE, M. E. (1994): *Images of mind*. New York: WH Freeman
- PULLEN, R. (1999): Psychotropic drugs and automobile driving ability of elderly patients. *Versicherungsmedizin*, 51, 71-74
- PUNTER, J. F. & KUBINGER, K. D. (2002): Was ist aus der Kritik der „Testrezensionen: 25 einschlägige Verfahren“ (*Zeitschrift für Differentielle und Diagnostische Psychologie*, 18, Heft 1-2) geworden?
- RATTEN, A., RATCLIFF, G., LONGDON, B., PETERS, B. & MORROW, L. (1990): Assessing the utility of the MVPT in predicting successful driving performance in head injured patients. *Archives of Clinical Neuropsychology*, 5, 205-295
- REGER, M. A., WELSH, R. K., WATSON, G. S., CHOLERTON, B., BAKER, L. D. & CRAFT, S. (2004): The relationship between neuropsychological functioning and driving ability in dementia: a meta-analysis. *Neuropsychology*, 18, 85-93
- REIMANN, G. (2005): Vom Rinnsal zum Strom: Die DIN 33430 bahnt sich ihren Weg. *Report Psychologie*, 03, 114-123
- RISSER, D., STICHENWIRTH, M., KLUPP, N., SCHNEIDER, B., STIMPFL, T., VYUDILIK, W. & BAUER, G. (1998): Drugs and driving in Vienna, Austria. *Journal of Forensic Sciences*, 43, 817-820
- RIZZO, M., REINACH, S., McGEHEE, D. & DAWSON, J. (1997): Simulated car crashes and crash predictors in drivers with Alzheimer disease. *Archives of Neurology*, 54, 545-551
- ROENKER, D. L., CISELL, G. M., BALL, K. K., WADLEY, V. G. & EDWARDS, J. D. (2003): Speed-of-processing and driving simulator training result in improved driving performance. *Human Factors*, 45, 218-233
- ROSSLER, H., BATTISTA, H. J., DEISENHAMMER, F., GUNTHER, V., POHL, P., PROKOP, L. & RIEMER, Y. (1993): Methadone-substitution and driving ability. *Forensic Science International*, 62, 63-66
- ROTHERMEL, S., SOMMER, S.-M. & De WAARD, D. (2003): The role of simulation in the training of novice drivers. In D. De WAARD, K.

- BROOKHUIS, S.-M. SOMMER & W. VERWEY (Hrsg.). Human Factors in the Age of Virtual Reality. Maastricht: Shaker Publishing
- RUBIN, G. S., NG, E. S., BANDEEN-ROCHE, K., KEYL, P. M., FREEMAN, E. E. & WEST, S. K. (2007): A prospective, population-based study of the role of visual impairment in motor vehicle crashes among older drivers: the SEE study. *Investigative Ophthalmology and Visual Science*, 48, 1483-1491
- SABATOWSKI, R., SCHWALEN, S., RETTIG, K., HERBERG, K. W., KASPER, S. M. & RADBRUCH, L. (2003): Driving ability under long-term treatment with transdermal fentanyl. *Journal of Pain and Symptom. Manage.*, 25, 38-47
- SCHLAG, B. (1993): Elderly drivers in Germany--fitness and driving behavior. *Accident Analysis and Prevention*, 25, 47-55
- SCHNEIDER, J. J., NOVACK, T. A., ALDERSON, A. & BUSH, B. A. (2000): UFOV risk level and driving ability in patients with traumatic brain injury. *Archives of Clinical Neuropsychology*, 15, 658-658
- SCHNEIDER, J. J. & GOUVIER, W. D. (2005): Utility of the UFOV Test With Mild Traumatic Brain Injury. *Applied Neuropsychology*, 12, 138-142
- SCHUBERT, W. & BERG, M. (2001): Zu einigen methodischen Fragen der Anwendung von psychologischen Testverfahren im Rahmen der Fahreignungsbegutachtung. *Zeitschrift für Verkehrssicherheit*, 1, 9-14
- SCHUBERT, W. & WAGNER, T. (2003): Die psychologische Fahrverhaltensbeobachtung – Grundlagen, Methodik und Anwendungsmöglichkeiten. *Zeitschrift für Verkehrssicherheit*, 3, 119-127
- SCHUBERT, W. & MATTERN, R. (2004): Urteilsbildung in der medizinisch-psychologischen Fahreignungsdiagnostik – Beurteilungskriterien. *Zeitschrift für Verkehrssicherheit*, 2, 89-96
- SCHUBERT, W., SCHNEIDER, W., EISENMENGER, W. & STEPHAN, E. (Hrsg.) (2005): Begutachtungs-Leitlinien zur Kraftfahrereignung – Kommentar. Bonn: Kirschbaum
- SCHULTE, T., STRASBURGER, H., MULLER-OEHRING, E. M., KASTEN, E. & SABEL, B. A. (1999): Automobile driving performance of brain-injured patients with visual field defects. *American Journal of Physical Medicine and Rehabilitation*, 78, 136-142
- SCHULTHEIS, M. T. & CHUTE, D. L. (1999): An ecologically valid driving assessment program: The neurocognitive driving test. *Archives of Clinical Neuropsychology*, 14, 794-795
- SCHULTHEIS, M. T. & MOURANT, R. R. (2001): Virtual reality and driving: the road to better assessment for cognitively impaired populations. *Presence*, 10, 431-439
- SCHUTZ, L. E. & SCHUTZ, J. A. (1996): Neuropsychological correlates of driving recovery after postacute rehabilitation. *Archives of Clinical Neuropsychology*, 11, 446-795
- SHALLICE, T. (1988): *From Neuropsychology to mental structure*. Cambridge: University Press
- SINGH, R., PENTLAND, B., HUNTER, J. & PROVAN, F. (2007): Parkinson's disease and driving ability. *Journal of Neurology, Neurosurgery and Psychiatry*, 78, 363-366
- SLATER, D. H. (2006): FAQs: neurological illness and driving. *Practitioner*, 250, 4-1
- SMUC, M., CHRIST, R. & GATSCHA, M. (2006): SAF (System zur Analyse des Fahrverhaltens)
- SODERSTROM, S. T., PETTERSSON, R. P. & LEPPERT, J. (2006): Prediction of driving ability after stroke and the effect of behind-the-wheel training. *Scandinavian Journal of Psychology*, 47, 419-429
- SOMERS, D. C., DALE, A. M., SEIFFERT, A. E. & TOOTELL, R. B. (1999): Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 1663-1668
- SOMMER, M., ARENDASY, M., OLBRICH, A. & SCHUHFRIED, G. (2004): Qualitätsverbesserung in der verkehrspsychologischen Diagnostik mit neuronalen Netzen: eine Pilotstudie. *Zeitschrift für Verkehrssicherheit*, 4, 193-199
- SOMMER, M. & HÄUSLER, J. (2006): Kriteriumsvalidität des Expertensystems Verkehr. *Zeitschrift für Verkehrssicherheit*, 2, 83-89

- SOMMER, S.-M. Humanist im Druck: Networking in the European Union for a Human-Centred Design of New Driver Support Systems. In: De WAARD, D., BROOKHUIS, K. A., Van EG-MOND, R. & BOERSEMA, T. (eds.): Human factors in design, safety, and management, 69-72. Maastricht: Shaker Publ., 2005
- SOMMER, S.-M., FALKMER, T., BEKIARIS, E. & PANU, M. (2004): Toward a Client-Centred Approach to Fitness-to-Drive Assessment in Elderly Drivers. *Swedish Journal for Occupational Therapy*, 11, 62-69
- SOYKA, M., HOCK, B., KAGERER, S., LEHNERT, R., LIMMER, C. & KUEFNER, H. (2005): Less impairment on one portion of a driving-relevant psychomotor battery in buprenorphine-maintained than in methadone-maintained patients: results of a randomized clinical trial. *Journal of Clinical Psychopharmacology*, 25, 490-493
- SPENCE, C. & DRIVER, J. (2004): Crossmodal space & crossmodal attention. Oxford: University Press
- SRIVASTAVA, A., MEHTA, R., ROTHKE, S. P., RA-DEMAKER, A. W. & BLEI, A. T. (1994): Fitness to drive in patients with cirrhosis and portal-systemic shunting: a pilot study evaluating driving performance. *Journal of Hepatology*, 21, 1023-1028
- STOLWYK, R. J., CHARLTON, J. L., TRIGGS, T. J., IANSEK, R. & BRADSHAW, J. L. (2006): Neuropsychological function and driving ability in people with Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, 28, 898-913
- STRUMPF, M., WILLWEBER-STRUMPF, A. & HERBERG, K. W. (2005): Sicherheitsrelevante Leistungsfähigkeit von Schmerzpatienten unter Langzeitopioidtherapie. *Schmerz*, 19, 426-433
- STURM, W., De, S. A., KRAUSE, B. J., SPECHT, K., HESSELMANN, V., RADERMACHER, I., HERZOG, H., TELLMANN, L., MULLER-GARTNER, H. W. & WILLMES, K. (1999): Functional anatomy of intrinsic alertness: evidence for a fronto-parietal-thalamic-brainstem network in the right hemisphere. *Neuropsychologia*, 37, 797-805
- STURM, W. & WILLMES, K. (2001): On the functional neuroanatomy of intrinsic and phasic alertness. *Neuroimage*, 14, S76-S84
- STURM, W. (2002): Diagnostics of Attention Deficits in Neurology. *Aktuelle Neurologie*, 29, 25-29
- STURM, W., LONGONI, F., FIMM, B., DIETRICH, T., WEIS, S., KEMNA, S., HERZOG, H. & WILLMES, K. (2004): Network for auditory intrinsic alertness: a PET study. *Neuropsychologia*, 42, 563-568
- Testkuratorium der Förderung deutscher Psychologenverbände (1986): Mitteilung. Beschreibung der einzelnen Kriterien für die Testbeurteilung. *Diagnostica*, 32, 358-360
- The British Psychological Society (1999): Fitness to drive and cognition: A Document of the Multi-Disciplinary Working Party on Acquired Neuropsychological Deficits and Fitness to Drive, British Psychological Society, Leicester, United Kingdom
- TSANG, P. S. & SHANER, T. L. (1998): Age, attention, expertise, and time-sharing performance. *Psychology and Aging*, 13, 323-347
- UC, E. Y., RIZZO, M., ANDERSON, S. W., SPARKS, J., RODNITZKY, R. L. & DAWSON, J. D. (2006): Impaired visual search in drivers with Parkinson's disease. *Annals of Neurology*, 60, 407-413
- UTTIL, B. & PILKENTON-TAYLOR, C. (2001): Letter cancellation performance across the adult life span. *Clinical Neuropsychology*, 15, 521-530
- Van ZOMEREN, A. H. & BROUWER, W. H. (1994): *Clinical Neuropsychology of attention*. New York: Oxford University Press
- VANCE, D. E., ROENKER, D. L., CISSELL, G. M., EDWARDS, J. D., WADLEY, V. G. & BALL, K. K. (2006): Predictors of driving exposure and avoidance in a field study of older drivers from the state of Maryland. *Accident Analysis and Prevention*, 38, 823-831
- VELDHUIJZEN, D. S., Van WIJCK, A. J. M., WILLE, F., VERSTER, J. C., KENEMANS, J. L., KALKMAN, C. J., OLIVIER, B. & VOLKERTS, E. R. (2006): Effect of chronic nonmalignant pain on highway driving performance. *Pain*, 122, 28-35
- VIVOLI, R., BERGOMI, M., ROVESTI, S., BUSSETTI, P. & GUAITOLI, G. M. (2006): Biological

- and behavioral factors affecting driving safety. *Journal of Preventive Medicine and Hygiene*, 47, 69-73
- Von MEYER, L. (2006): Opioids and driving ability. *MMW. Fortschritte der Medizin*, 148, 33-34
- WALTER, H., VETTER, S. C., GROTHE, J., WUNDERLICH, A. P., HAHN, S. & SPITZER, M. (2001): The neural correlates of driving. *Neuroreport*, 12, 1763-1767
- WEINGER, K., KINSLEY, B. T., LEVY, C. J., BAJAJ, M., SIMONSON, D. C., COX, D. J., RYAN, C. M. & JACOBSON, A. M. (1999): The perception of safe driving ability during hypoglycemia in patients with type 1 diabetes mellitus. *American Journal of Medicine*, 107, 246-253
- WHELIHAN, W., CARON, M., PIATT, A., OTT, B. & HEINDEL, W. (1998): Neuropsychological predictors of driving ability in early-stage dementia. *Archives of Clinical Neuropsychology*, 13, 50-795
- WHELIHAN, W. M., DICARLO, M. A. & PAUL, R. H. (2005): The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology*, 20, 217-228
- WITTKOWSKI, J. & SEITZ, W. (2006): Merkmale verkehrspsychologischer Eignungsgutachten der medizinisch-psychologischen Untersuchungsstellen. <http://psydok.sulb.uni-saarland.de/volltexte/2006/718/13.02.08>
- WITTMANN, M., KISS, M., GUGG, P., STEFFEN, A., FINK, M., POPPEL, E. & KAMIYA, H. (2006): Effects of display position of a visual in-vehicle task on simulated driving. *Applied Ergonomics*, 37, 187-199
- WOOD, J., CHAPARRO, A., HICKSON, L., THYER, N., CARTER, P., HANCOCK, J., HOE, A., LE, I., SAHETAPY, L. & YBARZABAL, F. (2006): The effect of auditory and visual distracters on the useful field of view: implications for the driving task. *Investigative Ophthalmology and Visual Science*, 47, 4646-4650
- WOOD, J. M. (2002): Age and visual impairment decrease driving performance as measured on a closed-road circuit. *Human Factors*, 44, 482-494
- ZIMMERMANN, P. & FIMM, B. (2002): Testbatterie zur Aufmerksamkeitsprüfung. Version 1.7. Herzogenrath: Psytest
- ZIMMERMANN, P. & FIMM, B. (2002): A test battery for attentional performance. In: M. LECLERCQ & P. ZIMMERMANN (eds.): *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation*. Hove: Psychology Press
- ZIMMERMANN, P. & LECLERCQ, M. (2002): Neuropsychological aspects of attentional functions and disturbances. In: M. LECLERCQ & P. ZIMMERMANN (eds.): *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation*. Hove: Psychology Press
- ZIMMERMANN, P. & FIMM, B. (2004): Die Testbatterie zur Aufmerksamkeitsprüfung (TAP). In: G. BÜTTNER & L. SCHMIDT-ATZERT (Hrsg.): *Diagnostik von Konzentration und Aufmerksamkeit*. Göttingen: Hogrefe

Schriftenreihe

Berichte der Bundesanstalt für Straßenwesen

Unterreihe „Mensch und Sicherheit“

2002

- M 135: **Nutzung von Inline-Skates im Straßenverkehr**
Alrutz, Gündel, Müller, Brückner, Gnielka, Lerner,
Meyhöfer € 16,00
- M 136: **Verkehrssicherheit von ausländischen Arbeitnehmern
und ihren Familien**
Funk, Wiedemann, Rehm, Wasilewski, Faßmann, Kabakci,
Dorsch, Klapproth, Ringleb, Schmidtpott € 20,00
- M 137: **Schwerpunkte des Unfallgeschehens von Motorradfahrern**
Assing € 15,00
- M 138: **Beteiligung, Verhalten und Sicherheit von Kindern und
Jugendlichen im Straßenverkehr**
Funk, Faßmann, Büschges, Wasilewski, Dorsch, Ehret, Klapproth,
May, Ringleb, Schießl, Wiedemann, Zimmermann € 25,50
- M 139: **Verkehrssicherheitsmaßnahmen für Kinder – Eine Sichtung
der Maßnahmenlandschaft**
Funk, Wiedemann, Büschges, Wasilewski, Klapproth,
Ringleb, Schießl € 17,00
- M 140: **Optimierung von Rettungseinsätzen – Praktische und
ökonomische Konsequenzen**
Schmiedel, Moecke, Behrendt € 33,50
- M 141: **Die Bedeutung des Rettungsdienstes bei Verkehrsunfällen
mit schädel-hirn-traumatisierten Kindern – Eine retrospektive Auswertung
von Notarzteinsetzprotokollen in Bayern**
Brandt, Sefrin € 12,50
- M 142: **Rettungsdienst im Großschadensfall**
Holle, Pohl-Meuthen € 15,50
- M 143: **Zweite Internationale Konferenz „Junge Fahrer und Fahrerinnen“**
€ 22,50
- M 144: **Internationale Erfahrungen mit neuen Ansätzen zur Absenkung
des Unfallrisikos junger Fahrer und Fahranfänger**
Willmes-Lenz € 12,00
- M 145: **Drogen im Straßenverkehr – Fahrsimulationstest, ärztliche
und toxikologische Untersuchung bei Cannabis und Amphetaminen**
Vollrath, Sachs, Babel, Krüger € 15,00
- M 146: **Standards der Geschwindigkeitsüberwachung im Verkehr
– Vergleich polizeilicher und kommunaler Überwachungsmaßnahmen**
Pfeiffer, Wiebusch-Wothge € 14,00
- M 147: **Leistungen des Rettungsdienstes 2000/01 – Zusammenstellung
von Infrastrukturdaten zum Rettungsdienst 2000 und Analyse des
Leistungsniveaus im Rettungsdienst für die Jahre 2000 und 2001**
Schmiedel, Behrendt € 15,00

2003

- M 148: **Moderne Verkehrssicherheitstechnologie – Fahrdaten-
speicher und Junge Fahrer**
Heinzmann, Schade € 13,50
- M 149: **Auswirkungen neuer Informationstechnologien auf das
Fahrverhalten**
Färber, Färber € 16,00
- M 150: **Benzodiazepine: Konzentrationen, Wirkprofile und Fahr-
tätigkeit**
Lutz, Stroheck-Kühner, Aderjan, Mattern € 25,50

- M 151: **Aggressionen im Straßenverkehr**
Maag, Krüger, Breuer, Benmimoun, Neunzig, Ehmanns € 20,00
- M 152: **Kongressbericht 2003 der Deutschen Gesellschaft für Ver-
kehrsmedizin e. V.** € 22,00
- M 153: **Grundlagen streckenbezogener Unfallanalysen auf Bun-
desautobahnen**
Pöppel-Decker, Schepers, Koßmann € 13,00
- M 154: **Begleitetes Fahren ab 17 – Vorschlag zu einem fahrpraxis-
bezogenen Maßnahmenansatz zur Verringerung des Unfallrisikos
junger Fahranfängerinnen und Fahranfänger in Deutschland**
Projektgruppe „Begleitetes Fahren“ € 12,50

2004

- M 155: **Prognosemöglichkeiten zur Wirkung von Verkehrssicher-
heitsmaßnahmen anhand des Verkehrszentralregisters**
Schade, Heinzmann € 17,50
- M 156: **Unfallgeschehen mit schweren Lkw über 12 t**
Assing € 14,00
- M 157: **Verkehrserziehung in der Sekundarstufe**
Weishaupt, Berger, Saul, Schimunek, Grimm, Pleßmann,
Zügenrucker € 17,50
- M 158: **Sehvermögen von Kraftfahrern und Lichtbedingungen im
nächtlichen Straßenverkehr**
Schmidt-Clausen, Freiding € 11,50
- M 159: **Risikogruppen im VZR als Basis für eine Prämiendif-
ferenzierung in der Kfz-Haftpflicht**
Heinzmann, Schade € 13,00
- M 160: **Risikoorientierte Prämiendifferenzierung in der Kfz-Haft-
pflichtversicherung – Erfahrungen und Perspektiven**
Ewers(t), Growitsch, Wein, Schwarze, Schwintowski € 15,50
- M 161: **Sicher fahren in Europa – 5. Symposium** € 19,00
- M 162: **Verkehrsteilnahme und -erleben im Straßenverkehr bei
Krankheit und Medikamenteneinnahme**
Holte, Albrecht € 13,50
- M 163: **Referenzdatenbank Rettungsdienst Deutschland**
Kill, Andrä-Welker € 13,50
- M 164: **Kinder im Straßenverkehr**
Funk, Wasilewski, Eilenberger, Zimmermann € 19,50

2005

- M 165: **Förderung der Verkehrssicherheit durch differenzierte An-
sprache junger Fahrerinnen und Fahrer**
Hoppe, Tekaas, Woltring € 18,50
- M 166: **Förderung des Helmtragens Rad fahrender Kinder und
Jugendlicher – Analyse der Einflussfaktoren der Fahrradhelmut-
zung und ihrer altersbezogenen Veränderung**
Schreckenberger, Schlittmeier, Ziesenis € 16,00
- M 167: **Fahrausbildung für Behinderte**
Zawatzky, Dorsch, Langfeldt, Lempp, Mischau € 19,00
- M 168: **Optimierung der Fahrerlaubnisprüfung – Ein Reformvor-
schlag für die theoretische Fahrerlaubnisprüfung**
Bönninger, Sturzbecher € 22,00
- M 169: **Risikoanalyse von Massenunfällen bei Nebel**
Debus, Heller, Wille, Dütschke, Normann, Placke,
Wallentowitz, Neunzig, Benmimoun € 17,00
- M 170: **Integratives Konzept zur Senkung der Unfallrate junger
Fahrerinnen und Fahrer – Evaluation des Modellversuchs im Land
Niedersachsen**
Stiensmeier-Pelster € 15,00
- M 171: **Kongressbericht 2005 der Deutschen Gesellschaft für
Verkehrsmedizin e. V. – 33. Jahrestagung** € 29,50
- M 172: **Das Unfallgeschehen bei Nacht**
Lerner, Albrecht, Evers € 17,50

- M 173: Kolloquium „Mobilitäts-/Verkehrserziehung in der Sekundarstufe“ € 15,00
- M 174: Verhaltensbezogene Ursachen schwerer Lkw-Unfälle
Evers, Auerbach € 13,50

2006

- M 175: Untersuchungen zur Entdeckung der Drogenfahrt in Deutschland
Iwersen-Bergmann, Kauert € 18,50
- M 176: Lokale Kinderverkehrssicherheitsmaßnahmen und -programme im europäischen Ausland
Funk, Faßmann, Zimmermann, unter Mitarbeit von Wasilewski, Eilenberger € 15,00
- M 177: Mobile Verkehrserziehung junger Fahranfänger
Krampe, Großmann € 15,50
- M 178: Fehlerhafte Nutzung von Kinderschutzsystemen in Pkw
Fastenmeier, Lehnig € 15,00
- M 179: Geschlechtsspezifische Interventionen in der Unfallprävention
Kleinert, Hartmann-Tews, Combrink, Allmer, Jüngling, Lobinger € 17,50
- M 180: Wirksamkeit des Ausbildungspraktikums für Fahrlehreranfänger
Friedrich, Brünken, Debus, Leutner, Müller € 17,00
- M 181: Rennspiele am Computer: Implikationen für die Verkehrssicherheitsarbeit – Zum Einfluss von Computerspielen mit Fahrzeugbezug auf das Fahrverhalten junger Fahrer
Vorderer, Klimmt € 23,00
- M 182: Cannabis und Verkehrssicherheit – Mangelnde Fahreignung nach Cannabiskonsum: Leistungsdefizite, psychologische Indikatoren und analytischer Nachweis
Müller, Topic, Huston, Strohbeck-Kühner, Lutz, Skopp, Aderjan € 23,50
- M 183: Hindernisse für grenzüberschreitende Rettungseinätze
Pohl-Meuthen, Schäfer, Gerigk, Moecke, Schlechtriemen € 17,50

2007

- M 184: Verkehrssicherheitsbotschaften für Senioren – Nutzung der Kommunikationspotenziale im allgemeinmedizinischen Behandlungsalltag
Kocherscheid, Rietz, Poppelreuter, Riest, Müller, Rudinger, Engin € 18,50
- M 185: 1st FERSI Scientific Road Safety Research-Conference
Dieser Bericht liegt nur in digitaler Form vor und kann kostenpflichtig unter www.nw-verlag.de heruntergeladen werden € 24,00
- M 186: Assessment of Road Safety Measures
Erstellt im Rahmen des EU-Projektes ROSEBUD (Road Safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making) € 16,00
- M 187: Fahrerlaubnisbesitz in Deutschland
Kalinowska, Kloas, Kuhfeld € 15,50
- M 188: Leistungen des Rettungsdienstes 2004/05 – Analyse des Leistungsniveaus im Rettungsdienst für die Jahre 2004 und 2005
Schmiedel, Behrendt € 15,50
- M 189: Verkehrssicherheitsberatung älterer Verkehrsteilnehmer – Handbuch für Ärzte
Henning € 15,00
- M 190: Potenziale zur Verringerung des Unfallgeschehens an Haltestellen des ÖPNV/ÖPSV
Baier, Benthaus, Klempf, Schäfer, Maier, Enke, Schüller € 16,00

- M 191: ADAC/BAST-Symposium "Sicher fahren in Europa" – Referate des Symposiums vom 13. Oktober 2006 in Baden-Baden
Dieser Bericht liegt nur in digitaler Form vor und kann kostenpflichtig unter www.nw-verlag.de heruntergeladen werden. € 24,00

2008

- M 192: Kinderunfallatlas
Neumann-Opitz, Bartz, Leipnitz € 14,50
- M 193: Alterstypisches Verkehrsrisiko
Schade, Heinzmann € 14,50
- M 194: Wirkungsanalyse und Bewertung der neuen Regelungen im Rahmen der Fahrerlaubnis auf Probe
Debus, Leutner, Brünken, Skottke, Biermann € 14,50
- M 195: Kongressbericht 2007 der Deutschen Gesellschaft für Verkehrsmedizin (DGVM e.V.) – zugleich 50-jähriges Jubiläum der Fachgesellschaft DGVM – 34. Jahrestag € 28,00
- M 196: Psychologische Rehabilitations- und Therapiemaßnahmen für verkehrsauffällige Kraftfahrer
Follmann, Heinrich, Corvo, Mühlensiep, Zimmermann, Klipp, Bornewasser, Glitsch, Dünkel € 18,50
- M 197: Aus- und Weiterbildung von Lkw- und Busfahrern zur Verbesserung der Verkehrssicherheit
Frühauf, Roth, Schyggulla € 15,50
- M 198: Fahreignung neurologischer Patienten – Untersuchung am Beispiel der hepatischen Enzephalopathie
Knoche € 15,00

2009

- M 199: Maßnahmen zur Verbesserung der visuellen Orientierungsleistung bei Fahranfängern
Müsseler, Debus, Huestegge, Anders, Skottke € 13,50
- M 200: Entwicklung der Anzahl Schwerstverletzter infolge von Straßenverkehrsunfällen in Deutschland
Lefering € 13,50
- M 201: Bedeutung der Fahrpraxis für den Kompetenzerwerb beim Fahrenlernen
Grattenthaler, Krüger, Schoch € 20,00
- M 202: Computergestützte Medien und Fahrsimulatoren in Fahrausbildung, Fahrerweiterbildung und Fahrerlaubnisprüfung
Weiß, Bannert, Petzoldt, Krens € 16,00
- M 203: Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung
Poschadel, Falkenstein, Pappachan, Poll, Willmes von Hinckeldey € 16,50

Alle Berichte sind zu beziehen beim:

Wirtschaftsverlag NW
Verlag für neue Wissenschaft GmbH
Postfach 10 11 10
D-27511 Bremerhaven
Telefon: (04 71) 9 45 44 - 0
Telefax: (04 71) 9 45 44 77
Email: vertrieb@nw-verlag.de
Internet: www.nw-verlag.de

Dort ist auch ein Kompletverzeichnis erhältlich.