# Traffic perception and hazard avoidance – Foundations and possibilities for implementation in novice driver preparation

Berichte der
Bundesanstalt für Straßenwesen

Mensch und Sicherheit Heft M 273b

# Traffic perception and hazard avoidance – Foundations and possibilities for implementation in novice driver preparation


by

Bianca Bredow
Roland Brünken
Annika Dressler
Tino Friedel
Jan Genschow,
Kristin Kaufmann
Sarah Malone
Susann Mörl
Mathias Rüdel
Thomas Schubert,
Dietmar Sturzbecher
Conrad Teichert
Winfried Wagner
Bernd Weiße



**Berichte der Bundesanstalt für Straßenwesen**

**Mensch und Sicherheit Heft M 273b**

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



Die Hefte der Schriftenreihe **Berichte der Bundesanstalt für Straßenwesen** können direkt bei der Carl Schünemann Verlag GmbH, Zweite Schlachtpforte 7, D-28195 Bremen, Telefon: (04 21) 3 69 03 - 53, bezogen werden.

Über die Forschungsergebnisse und ihre Veröffentlichungen wird in der Regel in Kurzform im Informationsdienst **Forschung kompakt** berichtet. Dieser Dienst wird kostenlos angeboten; Interessenten wenden sich bitte an die Bundesanstalt für Straßenwesen, Stabsstelle Presse und Öffentlichkeitsarbeit.

Ab dem Jahrgang 2003 stehen **die Berichte der Bundesanstalt für Straßenwesen** (BASt) zum Teil als kostenfreier Download im elektronischen BASt-Archiv ELBA zur Verfügung.
http://bast.opus.hbz-nrw.de

**Impressum**

**Verkehrswahrnehmung und Gefahrenvermeidung – Grundlagen und Umsetzungsmöglichkeiten in der Fahranfängervorbereitung.Innovationsbericht zum Fahrerlaubnisprüfungssystem 2011 – 2014. Hrsg.: TÜV | DEKRA arge tp 21**

**Autoren**
Bianca Bredow, Roland Brünken, Annika Dressler, Tino Friedel, Jan Genschow, Kristin Kaufmann, Sarah Malone, Susann Mörl, Mathias Rüdel, Thomas Schubert, Dietmar Sturzbecher, Conrad Teichert, Winfried Wagner, Bernd Weiße

**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**
Fachverlag NW in der
Carl Schünemann Verlag GmbH
Zweite Schlachtpforte 7, D-28195 Bremen
Telefon: (04 21) 3 69 03 - 53
Telefax: (04 21) 3 69 03 - 48
www.schuenemann-verlag.de

ISSN 0943-9315
ISBN 978-3-95606-333-6

Bergisch Gladbach, Juli 2017

# Abstract – Kurzfassung


**Traffic perception and hazard avoidance – Foundations and possibilities for implementation in novice driver preparation**

The present innovation report describes key topics addressed in the relevant research and development activities of the TÜV | DEKRA arge tp 21 working group during the report period from 2011 to 2014. Such essential aspects were (1) presentation of the scientific foundations for development and assessment of novice driver competence relating to traffic perception and hazard avoidance, (2) evaluation and further development of the theoretical driving test, (3) research and development work aimed at optimisation of the practical driving test, and (4) description of the continuing advances in vehicle engineering and their significance for both driver training and testing.

Re (1): Numerous empirical studies point to the competence deficits displayed by novice drivers compared to more experienced drivers in respect of traffic perception and hazard avoidance. The systematic development and acquisition of corresponding skills in advance of their solo participation in road traffic thus appears promising as a means to reduce novice-typical accident risks. The development of training and test concepts addressing traffic perception and hazard avoidance, however, requires prior description of the underlying psychological constructs. During the report period, the relevant perceptual and traffic psychology foundations were analysed alongside innovative, computer-based approaches which have already been implemented in international driving licence testing and experimental studies from the field of hazard perception research. A further study was conducted to survey teaching and learning offers for novice drivers. The results presented in the report provide a theoretical basis for the future anchoring of traffic perception and hazard avoidance in the system of novice driver preparation. This work must now be continued with the elaboration of innovative learning offers and corresponding assessment tasks for novice drivers.

Re (2): The work on the theoretical driving test referred to the evaluation of individual test items and their assignment to parallel tests, as well as further empirical studies to answer specifically arising questions (e. g. model calculations to enable the elaboration of empirically founded recommendations on optimisation of the test assessment). In 2014, i. e. at the end of the current report period, furthermore, test items based on dynamic situation presentations were used for the first time. The report provides an overview of the preparatory research and development work. Last but not least, initial results from the empirical testing of innovative task formats implementing new forms of response input (keyboard, foot pedal) and performance assessment (reaction time) are discussed. The research and development work outlined in the report illustrates how benefits of the new test medium “computer” have been utilised successfully for a number of years. It is equally evident, however, that significant further potential is embodied in computer-based testing – especially in respect of task development and test design; the objective for the future must be to exploit this potential.

Re (3): The research work on development of the practical driving test was essentially dominated by three projects: The BASt project “Optimisation of the Practical Driving Test” served to elaborate methodical foundations (it began in 2008 and was brought to conclusion during the first year of the report period). The second project was a feasibility study to investigate the practicability of an electronic test report, which was conducted by the TÜV | DEKRA arge tp 21 working group (2011 to 2012). The BASt project “Revision Project on an Optimised Practical Driving Test”, finally, was commenced in 2013 and is scheduled to continue beyond the current report period into 2015. Its purpose is to continue development of the corresponding contents, methods and procedures (including methods which will enable continuous evaluation). Once ready for actual use, the results are to be tested in selected model regions. Subsequent nationwide implementation of the elaborated standards would ensure that decisions on the eligibility to participate in motorised road traffic are based on a differentiated assessment of driving competence. At the same time, observed competence deficits could be recorded systematically and communicated to the driving licence applicant. This can be expected to strengthen the safety-enhancing impact of the practical driving test.

Re (4): Technical innovations in vehicle engineering – whether driver assistance systems, new drive concepts (e-mobility) or solutions for (partially) automated vehicle control – are entering the market at ever shorter intervals. Such developments bring changes in the demands placed on drivers and thus exert a decisive influence on the acquisition, assessment and upholding of driving competence. The report outlines the tasks resulting for the Technical Examination Centres in the future.

**Verkehrswahrnehmung und Gefahrenvermeidung – Grundlagen und Umsetzungsmöglichkeiten in der Fahranfängervorbereitung**

Der vorliegende Innovationsbericht beschreibt die Forschungs- und Entwicklungsschwerpunkte der TÜV | DEKRA arge tp 21 im Hinblick auf das Fahrerlaubnisprüfungssystem für den Zeitraum 2011-2014. Diese lagen (1) in der Aufbereitung wissenschaftlicher Grundlagen zur Vermittlung und Erfassung von Kompetenzen im Bereich der Verkehrswahrnehmung und Gefahrenvermeidung bei Fahranfängern, (2) in der Evaluation und Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung und (3) in Forschungs- und Entwicklungsarbeiten zur Optimierung der Praktischen Fahrerlaubnisprüfung sowie (4) in der Beschreibung des fahrzeugtechnischen Wandels und seiner Bedeutung für die Fahrausbildung und Fahrerlaubnisprüfung.

Zu (1): Kompetenzdefizite von Fahranfängern gegenüber erfahrenen Fahrern in der Verkehrswahrnehmung und Gefahrenvermeidung sind durch eine Vielzahl empirischer Studien belegt. Die systematische Vermittlung und Aneignung entsprechender Fähigkeiten im Vorfeld der selbstständigen Verkehrsteilnahme erscheint deshalb vielversprechend für die Verringerung fahranfängertypischer Unfallrisiken. Der Entwicklung diesbezüglicher Ausbildungs- und Prüfungskonzepte muss jedoch zunächst eine Beschreibung der zugrunde liegenden psychologischen Konstrukte vorausgehen. Im Berichtszeitraum wurden dazu die relevanten wahrnehmungs- und verkehrspsychologischen Grundlagen aufgearbeitet. Weiterhin wurden die international im Fahrerlaubniswesen bereits praktizierten innovativen PC-basierten Prüfungsansätze wie auch experimentelle Untersuchungen aus der „Hazard Perception“-Forschung analysiert. Schließlich wurden Lehr-Lernangebote für Fahranfänger recherchiert. Mit den im Bericht vorgelegten Ergebnissen wurde der theoretische Grundstein für die künftige Verankerung der Verkehrswahrnehmung und Gefahrenvermeidung in der Fahranfängervorbereitung gelegt. Hieran wird nun die Erarbeitung innovativer Lernangebote und Prüfungsaufgaben für Fahranfänger anschließen.

Zu (2): Die Arbeiten zur Theoretischen Fahrerlaubnisprüfung umfassten die Evaluation von Prüfungsaufgaben und Paralleltests sowie auch weiterführende empirische Untersuchungen zur Bearbeitung anlassbezogener Fragestellungen (u. a. Modellrechnungen zur Erarbeitung von empirisch gestützten Empfehlungen für eine optimierte Prüfungsbewertung). Weiterhin wurden 2014, d. h. zum Ende des Berichtszeitraums, erstmals Aufgaben mit dynamischen Situationsdarstellungen eingesetzt. Die vorbereitenden Forschungs- und Entwicklungsarbeiten werden im Bericht überblicksartig dargestellt. Nicht zuletzt werden erste Ergebnisse der empirischen Erprobung innovativer Aufgabenformate diskutiert, bei denen neuartige Formen der Antworteingabe (Tastendruck, Pedal) und der Leistungsbewertung (Reaktionszeit) verwendet wurden. Die skizzierten Forschungsarbeiten lassen erkennen, dass seit einigen Jahren die Vorteile des Prüfmediums „Computer“ erfolgreich genutzt werden. Sie zeigen jedoch auch, dass die PC-gestützte Prüfung weitere Potenziale – insbesondere im Bereich der Aufgabenentwicklung und Testkonstruktion – birgt, die es zukünftig auszuschöpfen gilt.

Zu (3): Die Arbeiten zur Praktischen Fahrerlaubnisprüfung waren im Wesentlichen durch drei verbundene Projekte gekennzeichnet: Das BASt-Projekt „Optimierung der Praktischen Fahrerlaubnisprüfung“ diente der Erarbeitung methodischer Grundlagen (es begann 2008 und endete im ersten Jahr des Berichtszeitraumes). Als zweites folgte eine von der TÜV | DEKRA arge tp 21 durchgeführte Machbarkeitsstudie (2011 bis 2012) zur Untersuchung der Praktikabilität einer elektronischen Prüfungsdokumentation (e-Prüfprotokoll). Das BASt-Projekt „Revision zu einer optimierten Praktischen Fahrerlaubnisprüfung“ stellt schließlich das dritte Projekt dar (es begann im Jahr 2013 und wird über den Berichtszeitraum hinaus bis 2015 fortgeführt). Es dient dazu, die Inhalte, Verfahren und Abläufe der Prüfung (einschließlich Verfahrensweisen zu einer kontinuierlichen Evaluation) bis hin zur Einsatzreife weiterzuentwickeln und in ausgewählten Modellregionen zu erproben. Mit einer bundesweiten Implementierung der erarbeiteten Standards würde gewährleistet, dass die Entscheidung über die Zulassung zum motorisierten Straßenverkehr künftig auf einer differenzierten Fahrkompetenzeinschätzung basiert und festgestellte Kompetenzdefizite systematisch erfasst und zurückgemeldet werden – dies lässt eine verbesserte Sicherheitswirksamkeit der Praktischen Fahrerlaubnisprüfung erwarten.

Zu (4): Fahrzeugtechnische Innovationen – seien es Fahrerassistenzsysteme, Antriebskonzepte (e-Mobilität) oder das (teil-)automatisierte Fahren – kommen in immer kürzeren Abständen auf den Markt. Diese Entwicklungen sind mit veränderten Anforderungen an das Fahren verbunden und haben daher maßgeblichen Einfluss auf den Erwerb, die Überprüfung und den Erhalt von Fahrkompetenz. Im Bericht werden die daraus resultierenden zukünftigen Aufgabenstellungen für die Technischen Prüfstellen skizziert.

# Kurzbericht


### Verkehrswahrnehmung und Gefahrenvermeidung – Grundlagen und Umsetzungsmöglichkeiten in der Fahranfängervorbereitung

Mit den Innovationsberichten wird die Fachöffentlichkeit über die Forschungs- und Entwicklungsergebnisse informiert, die mit der mittel- und langfristigen Weiterentwicklung des Fahrerlaubnisprüfungssystems zusammenhängen. Mithilfe der Innovationsberichte kann somit die Qualität, die Planmäßigkeit und die wissenschaftliche Absicherung der Weiterentwicklung der Fahrerlaubnisprüfung beurteilt werden. Der vorliegende Innovationsbericht beschreibt die Hauptschwerpunkte der Tätigkeit der TÜV | DEKRA arge tp 21 im Hinblick auf die Fahrerlaubnisprüfungen für den Berichtszeitraum 2011 – 2014. Diese lagen (1) in der Aufbereitung wissenschaftlicher Grundlagen zur Vermittlung und Erfassung von Kompetenzen im Bereich der Verkehrswahrnehmung und Gefahrenvermeidung bei Fahranfängern, (2) in der Evaluation und Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung und (3) in Forschungs- und Entwicklungsarbeiten zur Optimierung der Praktischen Fahrerlaubnisprüfung sowie (4) in der Beschreibung des fahrzeugtechnischen Wandels und seiner Bedeutung für die Fahrausbildung und Fahrerlaubnisprüfung.

Zu (1): Kompetenzdefizite von Fahranfängern gegenüber erfahrenen Fahrern im Bereich der Verkehrswahrnehmung und Gefahrenvermeidung sind durch eine Vielzahl empirischer Studien belegt. Die systematische Vermittlung und Aneignung entsprechender Fähigkeiten im Vorfeld der selbstständigen Verkehrsteilnahme erscheint deshalb vielversprechend für die Verringerung von fahranfängertypischen Unfallrisiken. Die zu bewältigenden Aufgaben für eine Verankerung der Verkehrswahrnehmung und Gefahrenvermeidung in der Fahranfängervorbereitung lassen sich entlang von vier Hauptschritten beschreiben:

- Beschreibung des psychologischen Konstrukts,
- Erarbeitung eines Ausbildungskonzepts,
- Erarbeitung eines Prüfungskonzepts,
- Implementierung in das Gesamtsystem der Fahranfängervorbereitung.

Implementierung in das Gesamtsystem der Fahranfängervorbereitung.Diese Schritte werden im vorliegenden Bericht durch die Aufbereitung des wissenschaftlichen Forschungsstandes sowie durch die Erarbeitung konzeptueller Grundlagen unterlegt.

Zur genaueren Beschreibung des psychologischen Konstrukts wurden verschiedene Ansätze zur verkehrspsychologischen Modellierung von Fähigkeiten im Bereich der Verkehrswahrnehmung und Gefahrenvermeidung vergleichend analysiert. Das Ergebnis dieser Modellanalyse ist eine differenzierte Beschreibung von zu bewältigenden Anforderungen bei der Verkehrswahrnehmung und Gefahrenvermeidung. Auf diesen theoretischen Rahmen wird bei der zukünftigen Entwicklung von Ausbildungs- und Prüfungskonzepten Bezug genommen.

Im Bericht werden des Weiteren Ergebnisse einer Recherche zu Ausbildungsangeboten mit Bezug zur Verkehrswahrnehmung und Gefahrenvermeidung in Deutschland sowie im internationalen Raum vorgestellt. Die Darstellung der Ausbildungsangebote beinhaltet jeweils neben einer Beschreibung ihrer methodisch-inhaltlichen Konzeptgrundlagen auch die Einordnung empirischer Wirksamkeitsbelege. Es bleibt festzuhalten, dass es in Deutschland – im Gegensatz zu anderen Ländern – bislang keine speziellen Ausbildungseinheiten zur Förderung von Kompetenzen zur Verkehrswahrnehmung implementiert sind. Mit dem Abschluss des BASt-Projekts „Ansätze zur Optimierung der Fahrschulausbildung in Deutschland" liegen jedoch entsprechende Konzepte vor (z. B. eine 90-minütige Ausbildungseinheit „Verkehrswahrnehmung und Gefahrenvermeidung im Straßenverkehr", die durch die Bundesvereinigung der Fahrlehrerverbände und das Institut für angewandte Familien-, Kindheits- und Jugendforschung an der Universität Potsdam entwickelt wurde); diese Ausbildungskonzepte werden derzeit empirisch erprobt.

Als Grundlage für die Erarbeitung eines zukünftigen innovativen Prüfungskonzepts wurden die international im Fahrerlaubniswesen bereits praktizierten PC-basierten Prüfungsansätze analysiert. Derzeit finden sich sog. „Hazard-Perception-Tests" weltweit in zwei europäischen Ländern (Niederlande, Großbritannien) sowie in fünf australischen Bundesstaaten. Im Hinblick auf die erfassten Leistungsparameter zeigt der Vergleich, dass im Wesentlichen die folgenden drei kognitiven Anforderungen bei der Aufgabenbearbeitung zu unterscheiden sind: das Erkennen von Situationsmerkmalen, die eine nicht näher bestimmte Handlung erfordern (Antizipationslatenz); das Erkennen von Situationsmerkmalen,

aus welchen sich die Auswahl einer spezifischen Handlung begründet (Handlungsauswahl, und das Erkennen von Situationsmerkmalen, die den sicheren Ausführungszeitpunkt einer spezifischen Handlung indizieren (Handlungszeitpunkt).

Über die Analyse von Prüfungsansätzen hinaus, wurden für den vorliegenden Innovationsbericht experimentelle Studien zur Verkehrswahrnehmung und Gefahrenvermeidung ermittelt und mit besonderem Blick auf die darin verwendeten Operationalisierungsansätze ausgewertet. Der Rückgriff auf diese Studien ermöglicht es, die zukünftige Entwicklung innovativer Aufgabenformate auf eine breite empirisch begründete Grundlage zu stellen. Die bislang erarbeiteten Operationalisierungsansätze sind dem vorliegenden Bericht in Form sog. Aufgabenkonzepte (d. h. einer Skizzierung von Prüfungsinhalten und Lehrzielen, des Instruktions- und Antwortformats, der Leistungsparameter und Be-Bewertungskriterien sowie Erwägungen zur testökonomischen und technischen Umsetzbarkeit) als Anlage beigefügt.

Die vorgelegten konkreten Umsetzungsvorschläge für innovative Prüfungsaufgaben für die Fahrerlaubnisprüfung können – ihre Bewährung in empirischen Erprobungsuntersuchungen vorausgesetzt – zukünftig im Rahmen eines computergestützten „Verkehrswahrnehmungstests“ eine eigenständige Prüfungsform neben der Theoretischen und Praktischen Fahrerlaubnisprüfung in Deutschland darstellen. Die begonnenen Arbeiten zur Entwicklung eines solchen Verkehrswahrnehmungstests werden im folgenden Berichtszeitraum fortgesetzt.

Zu (2): Die durchgeführten Arbeiten zur Evaluation und Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung (TFEP) umfassen zum einen die kontinuierliche Evaluation von Prüfungsaufgaben und Paralleltests sowie auch weiterführende empirische Untersuchungen zur Bearbeitung anlassbezogener Fragestellungen; Letztere betrafen beispielsweise Modellrechnungen zur Erarbeitung von empirisch gestützten Empfehlungen für eine Optimierung der Bewertungssystematik der Prüfung. Im Ergebnis dieser Untersuchung wird empfohlen, die derzeit mit 2 bis 5 Fehlerpunkten gewichtete Bewertungssystematik der TFEP durch eine einheitliche Punktebewertung für alle Aufgaben und die Verwendung eines „K.O.“-Kriteriums (d. h. beispielsweise keine zulässigen Falschbearbeitungen im Inhaltsbereich „Vorfahrt, Vorrang“) zu ersetzen. Die Anwendung dieser Bewertungssystematik lässt erwarten, dass die Unterscheidung von „Könnern“ und „Nichtkönnern“ aufgrund ihrer pädagogisch-didaktischen Begründung präziser erfolgt. Zugleich würde die Praktikabilität der Testkonstruktion wesentlich verbessert werden.

Nach der Einführung PC-generierter Abbildungen (anstelle von Fotos) für Bildaufgaben in der TFEP im Jahr 2011 stellen sog. „Abbildungsvarianten“ einen aufbauenden Entwicklungsschritt dar. Hierbei werden zu einer Bildaufgabe Varianten erzeugt, die sich hinsichtlich des erfassten Aufgabeninhalts (z. B. Lernziel, Auswahlantworten, Positionen von abgebildeten Verkehrsteilnehmern) nicht voneinander unterscheiden, jedoch in inhaltsirrelevanten Bildmerkmalen (z. B. Randbebauung der Verkehrsumgebung, Fahrzeugfarben, Fahrzeugtypen) Unterschiede aufweisen. Solche Varianten computergenerierter statischer Abbildungen werden mit dem Ziel entwickelt, eine stärkere inhaltliche Auseinandersetzung mit der bildlich dargebotenen Verkehrssituation in den traditionellen Mehrfach-Wahl-Aufgaben zu fördern und ein bloß oberflächliches Auswendiglernen von anzukreuzenden Auswahlantworten zu erschweren. Die Ergebnisse einer hierzu durchgeführten empirischen Untersuchung zeigen, dass die Variation von inhaltsirrelevanten Bildmerkmalen die Wiedererkennung erschwert und die Abbildungsvarianten – trotz Abweichung vom Originalbild – als inhaltlich-konzeptuell übereinstimmend wahrgenommen werden. Dies weist darauf hin, dass Aufgaben mit Abbildungsvarianten ihre intendierte prüfungsdidaktische Funktion erfüllen können.

Schließlich stellt der Einsatz von Mehrfach-Wahl-Aufgaben in Verbindung mit dynamischen Situationsdarstellungen im Jahr 2014, d. h. zum Ende des Berichtszeitraums, einen weiteren wichtigen Schritt zur Optimierung der TFEP dar. Die vorbereitenden Forschungs- und Entwicklungsarbeiten werden im vorliegenden Bericht überblicksartig dargestellt. Sie umfassten unter anderem die Festlegung geeigneter Rahmenbedingungen für die Einbindung dieser Aufgaben in die TFEP, die Erstellung einer hinreichenden Anzahl von entsprechenden Prüfungsaufgaben sowie die Durchführung von empirischen Erprobungsuntersuchungen. Bis zum Ende des Berichtszeitraums wurden 52 Aufgaben mit dynamischen Situationsdarstellungen in den amtlichen Fragenkatalog aufgenommen.

Weiterhin wurden wichtige Forschungsarbeiten zur empirischen Erprobung innovativer Antwortformate geleistet, die sich unter anderem auf das Medium der Antworteingabe (Tastendruck, Pedal) und die Parameter der Leistungsbewertung (Reaktionszeit) bezogen. Diesbezügliche empirische Quer- und Längsschnittuntersuchungen wurden am Lehrstuhl „Empirische Bildungsforschung" der Universität des Saarlandes durchgeführt. Die Ergebnisse erscheinen mit Blick auf die erprobten innovativen Reaktionszeitaufgaben mittels Tasteneingabe für die Weiterentwicklung der TFEP grundsätzlich vielversprechend. Die Längsschnittuntersuchungen zeigten, dass die Bearbeitungsleistungen im Lernverlauf zunehmen, also nachdem zu den angeeigneten Inhalten der theoretischen Ausbildung auch praktische Fahrerfahrungen hinzukommen.

Zu (3): Aufgrund der Optimierungsdynamik bei der Praktischen Fahrerlaubnisprüfung (PFEP) und der Überlegungen zur Weiterentwicklung des Gesamtsystems der Fahranfängervorbereitung empfahl die BASt im Jahr 2011, die Innovationsberichte zukünftig auf die PFEP zu erweitern. Dieser Empfehlung folgend, wird im vorliegenden Bericht zunächst dargestellt, welche Optimierungspotenziale für die PFEP bestehen und derzeit erschlossen werden. Darauf aufbauend werden die mittelfristigen Ziele bei der Implementierung einer optimierten PFEP beschrieben. Diese beinhalten nicht zuletzt, die Prüfungsleistungen der Bewerber detailliert zu erfassen und zurückzumelden sowie schließlich auch die testpsychologische Prüfungsgüte anhand von Evaluationsmechanismen zu beurteilen und zu verbessern.

Die Forschungs- und Entwicklungsarbeiten zur PFEP sind im Berichtszeitraum 2011 – 2014 im Wesentlichen durch drei Projekte gekennzeichnet: Das BASt-Projekt „Optimierung der Praktischen Fahrerlaubnisprüfung" begann 2008 und wurde im ersten Jahr des Berichtszeitraumes abgeschlossen. Im Rahmen der Projektarbeiten wurden Lösungsansätze zur Bewältigung der zentralen inhaltlichen und methodischen Herausforderungen für eine optimierte Fahrerlaubnisprüfung erarbeitet. So galt es zunächst, eine Anforderungskonzeption zu entwickeln, die den Prüfungsablauf durch die Festlegung besonders verkehrssicherheitsrelevanter prototypischer Beobachtungs- bzw. Verkehrssituationen (Fahraufgaben) mit hinreichend standardisierten Fahrverhaltensanforderungen (Fahrkompetenzbereiche) strukturieren und ihn steuerbar machen. Diese inhaltlich-methodischen Grundlagen sind die Voraussetzung für das Generieren aussagekräftiger (also objektiver, zuverlässiger und gültiger) Befunde zur Fahrkompetenz von Fahrerlaubnisbewerbern. Für die fachgerechte computergestützte Erfassung und elektronische Dokumentation der Prüfungsleistungen durch den Fahrerlaubnisprüfer im Prüfungsfahrzeug wurde eine methodische und technologische Konzeption (e-Prüfprotokoll) erarbeitet. Schließlich wird ein formatives und summatives Evaluationssystem beschrieben, auf dessen Grundlage die im Projekt erarbeiteten Durchführungs-, Anforderungs-, Bewertungs- und Dokumentationsstandards kontinuierlich weiterzuentwickeln sind.

Das zweite Projekt betrifft die Machbarkeitsstudie der TÜV | DEKRA arge tp 21 zum e-Prüfprotokoll, die von 2011 bis 2012 durchgeführt wurde. Die Bedeutung dieser Ersterprobung einer zukünftigen elektronischen Dokumentation der PFEP lag nicht allein darin zu prüfen, ob der Prototyp des Programms den funktionalen sowie soft- und hardwaretechnischen Anforderungen genügt; vielmehr wurde dabei der grundsätzlichen Frage nachgegangen, inwieweit eine elektronische Dokumentation der Prüfungsleistungen im Prüfungsfahrzeug überhaupt praktikabel ist und inwieweit sie für den Fahrerlaubnisprüfer ggf. eine Unterstützung im Arbeitsalltag darstellt. Im Ergebnis bleibt festzuhalten, dass sich die inhaltlichen und methodischen Grundlagen der PFEP als tragfähig für die Erarbeitung des elektronischen Prüfprotokolls erwiesen haben.

Das BASt-Projekt „Revision zu einer optimierten Praktischen Fahrerlaubnisprüfung" stellt schließlich das dritte Projekt dar; es begann im Jahr 2013 und wird über den Berichtszeitraum hinaus bis 2015 fortgeführt. Das Ziel des Revisionsprojekts besteht darin, die Inhalte, Verfahren und Abläufe der optimierten PFEP bis hin zur Einsatzreife weiterzuentwickeln, zu optimieren und zu erproben. Im Ergebnis soll ein tragfähiges methodisches Konzept für die computergestützte Durchführung sowie für die instrumentelle und prozessuale Evaluation der optimierten PFEP entstehen. Darüber hinaus dient dieses Projekt der Schaffung der technischen Voraussetzungen für die computergestützte Durchführung und Weiterentwicklung der optimierten PFEP. Den Kern des Projekts stellt die Praxiserprobung aller Verfahren und Abläufe in vier Modellregionen der mit der Fahrerlaubnisprüfung beliehenen Technischen Prüfstellen über einen Zeitraum von drei Monaten dar.

Zu (4): Fahrzeugtechnische Innovationen kommen in immer kürzeren Abständen auf den Markt und werden in vielen Variationen in Kraftfahrzeugen verbaut. Eindrucksvolle Beispiele dafür stellen zum einen Fahrerassistenzsysteme mit vernetzten Fahrzeugsicherheitsfunktionen und zum anderen neue Antriebskonzepte (e-Mobilität) dar. Auch erste Lösungen für das in Zukunft erwartete (teil-)automatisierte Fahren und der Einsatz neuartiger Fahrzeugkonzepte verdeutlichen das wachsende technische Innovationspotenzial. All diese Entwicklungen sind mit veränderten Anforderungen an das Fahren verbunden und haben daher maßgeblichen Einfluss auf den Erwerb, die Überprüfung und den Erhalt von Fahrkompetenz.

Will man die Anforderungen des absehbaren technischen Wandels an die Weiterentwicklung der Fahranfängervorbereitung ableiten, dann muss man zunächst prognostizieren, welche verkehrstechnischen Neuerungen in welchen Zeiträumen zu erwarten sind. So sind assistierende und teilautomatisierte Fahrfunktionen im Kraftfahrzeug bereits heute nutzbar. Der insbesondere mit Fahrerassistenzsystemen einhergehende Mehrnutzen im Komfortbereich und – wichtiger noch – im Bereich der Verkehrssicherheit kann nur ausgeschöpft werden, wenn die Fahrzeugführer die potenziellen Möglichkeiten wie auch Grenzen dieser Systeme kennen und ihre fachgerechte Bedienung in das gewohnte Fahrverhalten integrieren.

So ist beispielsweise mit Blick auf den Umgang mit Fahrerassistenzsystemen in der Fahrerlaubnisprüfung festzuhalten, dass die Fahrerlaubnisbewerber einerseits in der Lage sein müssen, die von ihnen genutzten Systeme sachgerecht und vorschriftsmäßig zu gebrauchen bzw. – sofern dies möglich ist, denn manche Fahrerassistenzsysteme kann man gar nicht ausschalten oder „bewusst“ handhaben − zu bedienen. Andererseits muss jedoch ebenso sichergestellt sein, dass die Bewerber auch ohne bestimmte Fahrerassistenzsysteme hinreichend sicher fahren können. Damit erscheint auch klar, dass die Fahrlehrer und Fahrerlaubnisprüfer das grundlegende Wissen und Können in Bezug auf die Nutzung bzw. Nichtnutzung von Fahrerassistenzsystemen sicher beherrschen müssen und dass es gewisse verkehrspädagogisch-didaktische und verkehrsrechtliche Vorgaben für die Verwendung von Fahrerassistenzsystemen in Ausbildung und Prüfung geben sollte, beispielsweise um eine einheitliche Fahranfängervorbereitung in Deutschland und nicht zuletzt die Prüfungsgerechtigkeit zu gewährleisten.

Die steigende Variantenvielfalt der Kraftfahrzeuge – zu denken ist hier beispielsweise an Pedelecs, S-Pedelecs, Trikes, Quades, Twizies oder Segways – stellt nicht nur eine Herausforderung für die Fahrerweiterbildung dar; vielmehr muss auch die Fahranfängervorbereitung auf dieses technische Phänomen reagieren. Eine in diesem Zusammenhang grundlegende Frage ist, ob die traditionelle Unterteilung der Fahrschulausbildung und Fahrerlaubnisprüfung nach Fahrerlaubnisklassen auch zukünftig sinnvoll ist bzw. ob technische Klassifikationen von Kraftfahrzeugen ggf. deckungsgleich mit fahrkompetenzbezogenen Klassifikationen von Fahrerlaubnissen sein müssen.

Neuartige Antriebsarten (z. B. elektrische Antriebe, verschiedene Hybrid-Varianten) lassen gegenüber „traditionellen Kraftfahrzeugarten“ Unterschiede bei der Fahrzeugbedienung erwarten. Hinzu kommt, dass funktionsgleiche Fahrzeugkomponenten bereits heute zunehmend herstellerspezifisch gestaltet werden und damit auch in unterschiedlicher Weise aktiviert und bedient werden müssen. Weitere neuartige Anforderungen bei der Fahrzeugbedienung ergeben sich z. B. aus der veränderten Leistungsentfaltung, den veränderten fahrphysikalischen Eigenschaften oder dem ungewohnten Bremsverhalten.

Durch den Einsatz von Kraftfahrzeugen mit hoch- bzw. vollautomatisierten Chauffeurfunktionen dürfte der zeitliche Anteil von Kontrollaufgaben bei der Führung des Kraftfahrzeugs wachsen, während der Anteil der Ausführung von Manövrieraufgaben sinkt. In Zukunft sind somit auch Veränderungen bei der Rolle des Menschen als Fahrzeugführer zu erwarten.

Mit Blick auf den kraftfahrzeugtechnischen Wandel und seine Implikationen für die Fahranfängervorbereitung werden im Bericht zu den vier Phänomenbereichen

- wachsende Verwendung von Fahrerassistenzsystemen,
- steigende Variantenvielfalt der Kraftfahrzeuge,
- Einführung neuer Antriebsarten und Fahrzeugkonzepte sowie
- Nutzung hochautomatisierter und vollautomatisierter Fahrfunktionen

kurze Überblicke über den technischen Entwicklungsstand und die zu erwartenden Entwicklungsprozesse gegeben. Dazu werden dann voraussichtliche Konsequenzen für die Weiterentwicklung der Fahranfängervorbereitung unter verkehrspädagogisch-prüfungsdidaktischen und verkehrsrechtlichen Gesichtspunkten prognostiziert sowie zukünftige Aufgabenstellungen der Technischen Prüfstellen abgeleitet.

# Contents

*Dietmar Sturzbecher, Mathias Rüdel & Jan Genschow*

# Innovation reports as a means for the further development of driving licence testing

## 1.1 The theoretical and practical driving tests in the system of novice driver preparation

The system of novice driver preparation in Germany (see Fig. 1) is characterised by a diversity of measures which are intended to offer novice drivers opportunities to participate in motorised road traffic, while at the same time reducing the associated risks. The model "Accompanied driving from age 17" ("Begleitetes Fahren ab 17", "BF17") and measures such as the introduction of a "zero alcohol" rule for novice drivers during the probationary period are prominent examples for the fundamental transformation of the system over the past few years. Essential system components – alongside professional driving school training – are the theoretical and practical driving tests (hereafter also referred to as "TDT" and "PDT", respectively). Both of these tests can be viewed as key measures for assessment of a novice driver's competence. They serve to define minimum training standards, which in turn describe the knowledge and skills required for participation in motorised road traffic, and thus the core of the training content for driving school training (Sturzbecher, Mörl & Kaltenbaek, 2014).

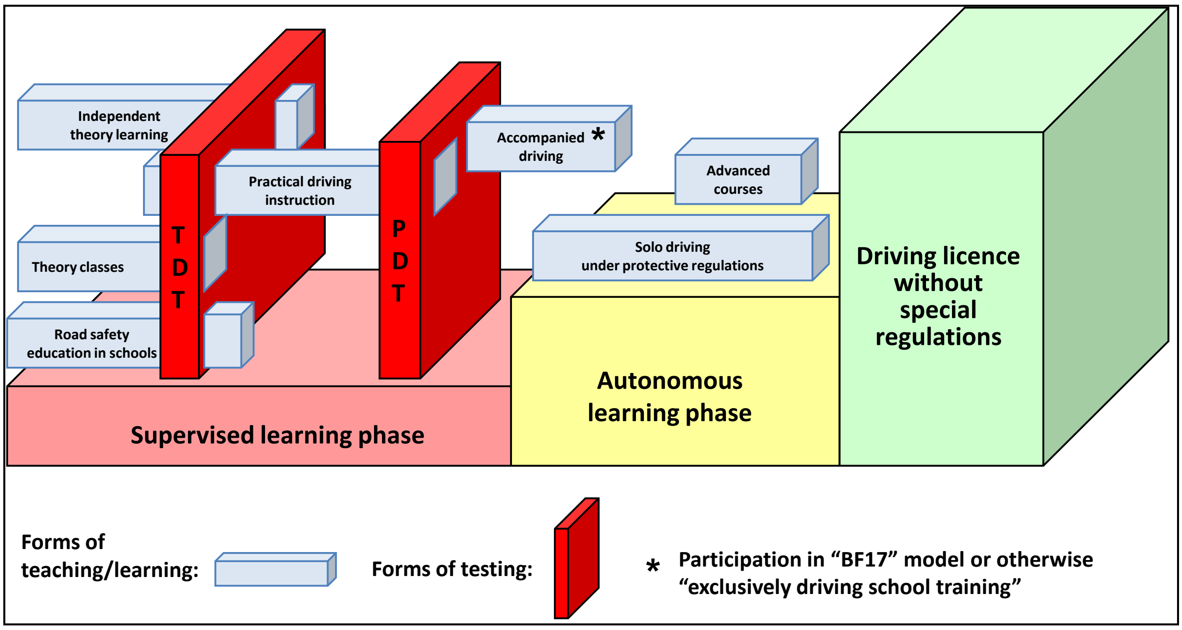



**Fig. 1: System of novice driver preparation in Germany**

The theoretical driving test takes the form of a "knowledge test": It serves to assess traffic-related explicit[1] knowledge (e.g. knowledge of traffic rules or appropriate behaviour in respect of traffic perception and hazard avoidance), and is presented to the candidate in the

[1] In the context of long-term knowledge, "explicit knowledge" is understood to mean factual or declarative knowledge. This comprises semantic or abstract knowledge of concepts, objects, facts, conditions or rules, as well as episodic or situated knowledge, in the form of situation prototypes and action scripts serving as central elements of top-down action planning. The descriptor "explicit" indicates that this knowledge can generally be reported and thus also conveyed and tested by way of verbal instruction or questioning (Grattenthaler, Krüger & Schoch, 2009).

form of text-based statements – and increasingly also by way of instruction formats with dynamic illustration of a given situation. The candidate is required to specify whether the given statements are true or false. The practical driving test, on the other hand, is from the psychological perspective a “work sample” which the test candidate completes within the framework of a test drive in real traffic. The candidate is here expected to furnish proof that he[2] is in a position to master the demands of traffic participation with adequate confidence, for which it is necessary to demonstrate a certain minimum level of so-called “process knowledge”[3] in respect of correct action execution.

As can be seen from Fig. 1, the TDT and PDT are usually positioned at different stages of the overall system of novice driver preparation, which in turn results in specific relationships to other system components. The TDT, for example, is taken after the completion of theory classes in the driving school and usually before or shortly after commencing practical driving instruction. To date, therefore, practical driving experience on the part of the novice driver is not an imperative prerequisite for successful completion of the TDT. The PDT, by contrast, is not taken until the end of practical driving instruction, i.e. only after the driving instructor has determined that the candidate is “ready for the test”. At this stage of training, the candidate is already able to retrieve and evaluate targeted information from complex traffic situations, and subsequently to translate this information into behaviour appropriate to the given situation. Passing of the PDT immediately broadens the scope of the successful candidate's driver rights: Participants in the “BF17” model become entitled to drive independently if accompanied by an experienced driving licence holder (subject to having reached the required minimum age); under the model with exclusively driving school training, successful completion of the test already marks the transition from the “supervised learning phase” to solo driving without obligations in respect of supervising passengers (“autonomous learning phase”).

Despite the different positioning within the process of driving competence acquisition, and similarly the different methodical approaches to assessment of this competence, it is common to both forms of driver testing – the TDT as a computer-based knowledge test and the PDT as a driving test in real traffic – that they aim to demonstrate the qualification to drive a motor vehicle, as specified by § 15 (1) of the Driving Licence Regulations (Fahrerlaubnisverordnung, FeV). In accordance with the applicable legislation, specifically § 2 (5) of the Road Traffic Act (Straßenverkehrsgesetz, StVG), a person is deemed appropriately qualified “if he or she

1. has acquired adequate knowledge of the legal regulations decisive for the driving of a motor vehicle,
2. is acquainted with the hazards of road traffic and with the behaviour necessary to avert such hazards,
3. possesses the technical knowledge required to operate a motor vehicle safely, where appropriate together with a corresponding trailer, and is capable of applying this knowledge in practice, and
4. possesses sufficient knowledge of an environment-aware and energy-saving manner of driving and is capable of applying this knowledge in practice.”

[2] Wherever gender-specific nouns or pronouns are used in this report, this serves solely to maximise general legibility and is in all cases to be understood to refer to persons of both genders.

[3] Process knowledge integrates explicit and implicit knowledge (i.e. procedural components of long-term knowledge, which are acquired in the form of motor schemata and developed through practice and the gathering of experience under changing action conditions), the product of which is a corresponding scope of (automated) psychomotor skills (Grattenthaler et al., 2009).

For strategic orientation of the continuous further development of driving licence testing, it thus seems expedient to pursue a joint analysis of both forms of testing, so as to identify their specific methodical potentials and limitations for the assessment of novice driver competence. This understanding of driving licence testing, with different components to evaluate corresponding partial competences, is already influencing current work to develop innovative test approaches and in this sense provides a foundation for the further discussions in the present innovation report.

## 1.2 Function and purpose of the innovation reports

Documentation of the further development of driving licence testing[4] in the manner described in the "System Manual on Driver Licensing (Theory Test)" ("Handbuch zum Fahrerlaubnisprüfungssystem (Theorie)", TÜV | DEKRA arge tp 21, 2008) was realised for the first time with presentation of an innovation report covering the period 2009/2010. This form of recurrent reporting on the course of evaluation and further development processes is an essential element of the optimised driving licence testing system. In its comment on the "Innovation Report on Optimisation of the Theoretical Driving Test – Report Period 2009/2010" (TÜV | DEKRA arge tp 21, 2011), the Federal Highway Research Institute (Bundesanstalt für Straßenwesen, BASt) drew attention to the dynamic progress of work on the optimisation of practical testing and – taking up thoughts on further development of the system of novice driver preparation as a whole – recommended that future innovation reports should be expanded to embrace also the practical driving test. This recommendation has been followed in the present innovation report. To accommodate this thematic broadening, the federal/regional expert committee BLFA-FE/FL gave its approval for extension of the regular report period on 19th September 2013. The present report thus now covers research and development work conducted in the field of driving licence testing over the years 2011 to 2014.

The innovation reports document all significant research and development processes in connection with medium- and long-term further development of the driving licence testing system, and communicate the results attained to the relevant authorities. In accordance with the "System Manual on Driver Licensing (Theory Test)"[5] and the draft for a corresponding "System Manual on Driver Licensing (Practical Test)"[6], they should include the following content:

- Test-relevant study results on the current state of research in the traffic sciences (e.g. accident analysis) and the pertinent basic disciplines (e.g. engineering sciences, psychology, education and medicine), as well as on associated topics applicable

[4] Further forms of documentation within the framework of an optimised driving licence testing system are applications for the official approval of test item prototypes, in which the case-specific elaboration of individual prototypes for test items is described, and the annual progress reports in which the essential work processes in connection with the implementation, evaluation and maintenance of the driving licence testing system are recorded.

[5] This manual received the approval of the Federal/Regional Expert Committee "Driver Licensing and Driving Instructor Legislation" (Bund-Länder-Fachausschuss "Fahrerlaubnisrecht/Fahrlehrerrecht", BLFA-FE/FL) on 6th November 2008 and, at the behest of the federal ministry responsible for road traffic, has since served as a common basis for the realisation and further development of the theoretical driving test.

[6] This manual has been presented in draft form and outlines processes and procedures for a future, optimised practical driving test, together with measures relating to continuous quality assurance and further development (Sturzbecher, Mörl & Kaltenbaek, 2014).

to the system of novice driver preparation (e.g. new forms of training and test content, or innovative teaching and learning tools),

- The results of own research work, especially on the development of improved test item formats serving optimisation of the TDT (e.g. new styles of content and innovative methodical formats, new test methods) and improved demand and assessment standards aimed at optimisation of the PDT (e.g. new test elements, driving tasks, observation categories and assessment criteria), alongside possibilities for optimised models of novice driver preparation comprising innovative combinations of further developed training and test elements,
- Information on the planning and results of work to establish the necessary prerequisites for optimisation of the driving licence testing system (e.g. work in respect of technical further development, new scientific methods and research strategies, including special model and development projects),
- Pointers to any possibly necessary amendment of the legislation and guidelines relating to driver licensing.

The innovation reports are thus intended, above all, to provide scientific background information on the research and development work performed by the TÜV | DEKRA arge tp 21 working group. Furthermore, they are to present the strategic alignment and methodology of this work, along with the most important results and the derived consequences and planning for the system of driving licence testing. In this way, the innovation reports permit assessment of the progress and the scientific basis for the further development of driving licence testing. Last but not least, the innovation reports serve to demonstrate that the further development of the test is in line with advances in the traffic sciences and with the status of developments in the overall system of novice driver preparation.

## 1.3 Starting positions for elaboration of the innovation report for the period 2011 to 2014

The BASt project "Framework concept for the further development of novice driver preparation in Germany (AP F1100.4409017)" implemented a strategic realignment of work serving the improvement of novice driver safety. Through the transition to a comprehensive and scientifically founded optimisation strategy taking into account all relevant potential measures, the safe preparation of novice drivers for their independent participation in motorised road traffic is to be further improved in the future. Supported by a group of external scientific experts[7] and with contributions from representatives of both professional practice (driving instructors, Technical Examination Centres, road safety associations) and the federal and state governments, the BASt elaborated a framework concept to identify correspondingly founded design perspectives and to evaluate the safety-enhancing potential of the individual measures of novice driver preparation.

The framework concept attaches particular safety relevance to the driving licence tests and derives this significance from the specific functions they fulfil: On the one hand, they serve to ensure that only novice drivers with adequate qualification are entitled to participate in motorised road traffic ("selection function"); on the other hand, the test demands provide orientation for the processes of novice driver preparation by defining minimum standards

[7] Members of the expert group: MR (retd.) Dieter Hartmann, Hildesheim, Prof. Dr. Hans-Peter Krüger, University of Würzburg, Prof. Dr. Detlev Leutner, University of Duisburg-Essen, Prof. Dr. Dietmar Sturzbecher, Institute for Applied Research on Childhood, Youth and the Family (IFK) at the University of Potsdam

for learning objectives and content ("control function"). To further improve the effectiveness of these functions, the following points were recommended:

(1) Introduction of innovative test items to enable assessment of the candidate's abilities in terms of traffic perception and hazard avoidance,

(2) Enhancement of the methodical quality of the PDT through a process of continuous evaluation and the use of an electronic test report,

(3) Fundamental reappraisal of the test formats used to assess traffic and driving competence and the timing of such assessments within the overall course of novice driver preparation (BASt-Expertengruppe, 2012).[8]

The starting positions for implementation of these recommendations, as they existed at the beginning of the report period 2011-2014, are outlined in the following sections.

re (1): The use of a computer as test medium facilitates not only continuous evaluation of the TDT, but also the development, testing and implementation of innovative instruction and answer formats. At the beginning of 2011, all image-based test items used in the TDT were still presented in the form of conventional photographs or simple graphic illustrations of traffic situations[9], though research and development work aimed at utilising the potential of the computer for the visualisation of traffic situations had already commenced in various forms (cf. Innovation Report 2009-2010; TÜV | DEKRA arge tp 21, 2011). This included the creation of computer-generated static images as reproductions of the conventional test item illustrations (photographs, simple graphics) specified in the official catalogue of test questions, as well as the elaboration of different image variants for one and the same test item. Prototype drafts similarly existed for multiple-choice questions wherein a static image was replaced in the task instructions with an approximately 15-second dynamic scenario to illustrate a given situation. Within the framework of the BASt project "Test psychology and learning theory underpinning the test questions in theoretical driving tests, with special consideration of test question formats with image sequences (FE 82.326/2007)", furthermore, work was conducted to develop and test innovative question formats in which the candidate is allowed only a limited time for the response to a static image or dynamic scenario (Malone, Biermann, Brünken & Buch, 2012). Alongside, the results of the BASt project "Novice driver preparation in Europe (FE 82.325/2007)" presented a systematic analysis of novice driver preparation in a total of 44 countries and provided various new insights into the diverse forms of testing (e.g. "knowledge test", "driving test", "traffic perception test") in use in international practice (Genschow, Sturzbecher & Willmes-Lenz, 2013). It was at that same time, finally, that the potential of a computer-based "traffic perception test"[10] focussed on the assessment of competence pertaining to traffic perception and hazard avoidance was introduced into the scientific discussion as an

---

[8] These recommendations were derived directly from the research and development work conducted on optimisation of the theoretical and practical driving tests over the past ten years. This can be taken as indication that a systematic and scientifically founded process of optimisation and further development, as demanded by the BASt and the experts involved in the project for all measures relating to novice driver preparation, already exists in respect of driving licence testing.

[9] These illustrations were replaced with computer-generated images from 01.07.2011.

[10] In the English-language literature, the expression "hazard perception test" is often found as a generic term describing the most varied test approaches and demands relating to the detection of hazards arising in road traffic. In Great Britain, however, "Hazard Perception Test" refers to a specific test procedure (see Chapter 2.3.2). In the present report, therefore, the somewhat broader term "traffic perception test" is preferred, firstly because it does not imply any specific test procedure, and secondly as it also seems more precise from the didactic perspective (see Chapter 2.2.3).

additional form of testing alongside the TDT as a knowledge test and the PDT as a driving test.

re (2): As the outcome of a project initiated and financed by the Technical Examination Centres, a research report on the methodical foundations and possibilities for further development of the practical driving test was published in 2008 (Sturzbecher, Bönninger & Rüdel, 2008). In this report, the PDT was described from the test psychology perspective as a “work sample” assessed by way of systematic behaviour observation. Accordingly, the criterion-referenced assessment of a novice driver's performance takes place in the “life-world domain” of motorised road traffic, in which the test demands can only be standardised to a limited extent. The driving test examiner must thus apply an adaptive test strategy, which permits a founded test decision to be reached via a chosen succession of individual action steps. The necessary basis, i.e. demand standards in the form of driving tasks, observation categories, and assessment and decision criteria, however, was at that time still lacking adequate scientific foundation and structure; the relevant standards had to date evolved historically and were in essence still anchored in different legal provisions, such as the Driving Licence Regulations (FeV) and Examination Guidelines (Prüfungsrichtlinie). To enable further development of the PDT, therefore, the authors recommended modernisation of the existing driving tasks, in conjunction with restructuring and streamlining of the criteria applied to assess the performance of a driving test candidate, which, in their current form, adhered above all to the wording of the legal stipulations. The integration of defined driving tasks, observation categories and assessment criteria into an innovative style of test report (electronic test report), which could be made available and used on a mobile device carried by the driving test examiner, was intended to permit differentiated assessments and corresponding feedback to the test candidate, as well as facilitating continuous evaluation and systematic further development of the PDT as such. The implementation of these recommendations and continuation of the conceptual development work followed in a subsequent BASt project “Optimisation of the practical driving test (FE 82.0345/2008)”, in which a contextual and methodical (implementation) concept for continuous maintenance, quality assurance and further development of the PDT was elaborated (Sturzbecher, Mörl & Kaltenbaek, 2014).

re (3): The outlined developments relating to the theoretical and practical driving tests will in future yield additional and to a certain extent novel possibilities for the assessment of novice driver competence. At the same time, measures such as the accompanied driving model “BF17” and advances in the field of computer-assisted learning are contributing to diversification of the possibilities for driving competence acquisition. As far as driving licence testing is concerned, these system changes raise questions as to the appropriate positioning of the individual tests. These questions were also taken up in the BASt project “Approaches to the optimisation of driver education in Germany (FE 82.515/2010)” at the beginning of 2012, and thereby linked with the further development of driving school training in general and reformation of its curricula and foundations in teaching/learning theory in particular (Bredow & Sturzbecher, in press). On the part of the Technical Examination Centres, work continued on international comparative analyses of the different test models in practical use and on a more precise description and modelling of “driving competence” as the construct to be measured in driving licence tests (TÜV | DEKRA arge tp 21, 2011).

## 1.4 Objectives of the present innovation report

The “Innovation Report on Optimisation of the Theoretical Driving Test” (TÜV | DEKRA arge tp 21, 2011) was presented in 2011. The key tasks addressed during the documented report period, namely the years 2009 and 2010, were (1) the elaboration of conceptual foundations for a driving competence model, (2) evaluation and further development of the traditional question formats and the corresponding test methodology, (3) research serving the development of computer-generated dynamic driving scenarios as new instruction formats for the TDT, and (4) research on innovative test items suitable for the assessment of those dimensions of competence which relate to traffic perception and hazard avoidance. These aspects also represented essential content of the topics treated during the period of the present report (2011 to 2014): On the one hand, both the work on continuous evaluation of the TDT and initial research and development efforts in the field of computer-generated traffic situation visualisation were continued. At the same time, further systematic consideration of those methodical and competence theory notions which could be deemed applicable for the assessment of novice driver competence relating to traffic perception and hazard avoidance have in the meantime yielded a sound foundation for evaluation of the potential of "traffic perception tests" as a means to improve the system of novice driver preparation in Germany.

As explained at the beginning, the innovation reports are intended to describe and evaluate the different forms of testing which are available for the assessment of novice driver competence. Accordingly, the present innovation report describes research and development work conducted on optimisation of the theoretical and practical driving tests over the years 2011 to 2014:

- Chapter 2 offers firstly a differentiated description of the competence deficits displayed by novice drivers in respect of traffic perception, alongside the (partial) competences which are to be assessed within the framework of a driving licence test. To this end, relevant principles from the field of cognitive psychology are presented, possible methodical approaches (and here, in particular, those from other countries) are discussed, and results and research designs from experimental basic research are evaluated. Finally, an analysis of differences in the way testing is arranged in various international systems of novice driver preparation serves as a basis for a discussion of meaningful solutions with regard to the positioning and combination of different forms of testing in Germany.
- The international perspective is maintained in Chapter 3, which presents concepts for the development of skills relating to traffic perception and hazard avoidance, and discusses currently available empirical findings on the feasibility of training such skills. To achieve a structured description of the training concepts, references to different forms of teaching and learning applied in novice driver preparation (e.g. theory classes, practical driving instruction) are developed. Furthermore, concrete research and development projects relating to driving school training in Germany are presented, including work aimed at supplementing traditional forms of teaching and learning with instruments to enable computer-assisted competence acquisition relating to traffic perception and hazard avoidance.
- Chapter 4 is devoted to work on continuous evaluation and further development of the TDT during the report period from 2011 to 2014. Given that the previous innovation report already presented a detailed summary of the measures implemented to optimise test methodology since the introduction of a computer-based theoretical test, the content here refers above all to innovations in terms of the "instruction formats" (e.g. the introduction of computer-generated images in the TDT) and the associated research work. While important steps have already been undertaken in connection with the instruction formats, the first design ideas for innovative "answer formats" – especially with regard to the assessment of competence relating to traffic perception and hazard avoidance – are currently still being developed and tested. This research work is also to be presented in the present report.

- The use of a computer as test medium offers improved possibilities for the visualisation of traffic environments and for realistic presentations of dynamic situations, but in turn poses new questions regarding the validity of test contents. Where tests depend increasingly on demonstration of the action knowledge required to master certain traffic situations (here presented in the form of virtual reality sequences), questions relating to a scientifically founded selection of appropriate situations and demands for the individual test items are shifted into the foreground. Chapter 5 presents corresponding possibilities for systematisation, with which the diversity of possible traffic situations can be structured and finally also an evaluation of the importance of different situational demands for driving licence testing is facilitated.
- Following the BASt recommendation to extend the contents of future innovation reports to cover also the PDT, Chapter 6 first describes the potential for optimisation of the PDT and those aspects of this potential which are currently being developed. On this basis, the medium-term objectives for implementation of an optimised PDT in Germany are described. This includes not least detailed documentation and feedback on the candidate's test performance, as well as evaluation mechanisms by which to assess and enhance the quality of the test as such from the perspective of psychological testing.
- Parallel to the developments in test methodology, significant advances are to be witnessed in vehicle engineering and technology. These advances are manifested, for example, in a diversity of electronic information, communication and control systems serving to support drivers (e.g. driver assistance systems, tutorial systems), in new power train concepts (“e-mobility”) and in (partially) automated driving functions. Furthermore, vehicles are being developed which no longer fit the picture of what is today considered a “normal” motor vehicle and thus point to a pending pluralisation of vehicle concepts. Innovations in motor vehicle technology also hold potential to enhance traffic safety, provided novice drivers acquire knowledge of the possibilities and limitations of such systems and concepts within the framework of driver training, and provided this knowledge is verified in a subsequent driving licence test. It is thus necessary for the test content to take account of such innovations. Their use during the driving test, moreover, may influence competence assessment. In Chapter 7, therefore, a number of technical developments are presented as examples, and the resulting demands in the context of driving licence testing are outlined.

*Jan Genschow & Dietmar Sturzbecher*

# Traffic perception tests as an innovative form of testing in novice driver preparation

## 1.5 Overview

Work on optimisation of the system of driving licence testing in Germany since 1999 has followed two separate lines of research and development, one geared to the theoretical driving test, the other to the practical driving test. The common starting point was the study project "Optimisation of driving licence testing (FE 82.113/1997)", which was sponsored by the Federal Highway Research Institute (BASt) to promote the scientific further development of content-related and methodical foundations for driving licence testing. The project report concentrates on proposals as to how the test contents for the TDT could be extended, and how the methodical quality of its assessment and evaluation could be enhanced by using a computer as test medium (Bönninger & Sturzbecher, 2005). Continuous optimisation of the PDT began in 2005 with an independent research project initiated by the Technical Examination Centres and is reviewed in more detail in Chapter 6.

Significant progress has been made on optimisation of the TDT since the initial presentation of the aforementioned research report (see Fig. 2). First of all, mechanisms for the continuous evaluation of parallel tests and individual test items were elaborated and tested. These mechanisms were subsequently anchored in a "System Manual on Driver Licensing (Theory Test)" (TÜV | DEKRA arge tp 21, 2008) in 2008. With the nationwide introduction of a computer-based TDT in 2010, finally, it was also possible to make use of the potential embodied in the new test medium with regard to improvements in test design and content visualisation. The randomised ordering of test items in the parallel tests, and likewise of the possible answers to a particular test item, for example, served to reduce the susceptibility to manipulation.

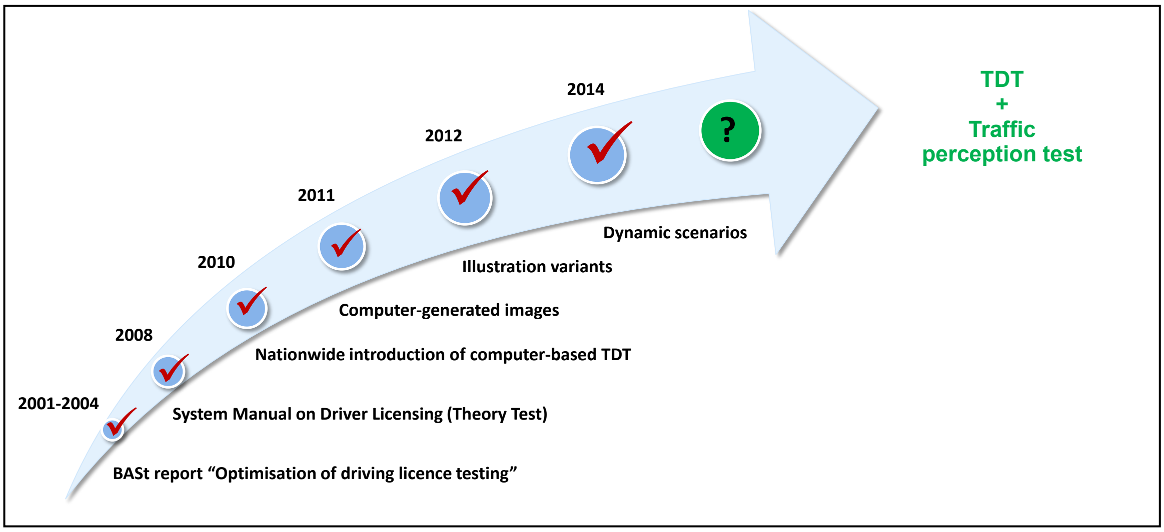



**Fig. 2: Milestones in the optimisation of the TDT**

In 2011, the photographs which had been used in illustrated test items in the past were replaced with computer-generated images. This was followed a year later by the first image-based test items with so-called "illustration variants", in which certain aspects of the depicted traffic situation (e.g. road users, traffic environment) were varied in their graphic presentation, without altering the fundamental item content. This was intended to inhibit the superficial memorisation of answer patterns and to force the test candidate to address the underlying item content more consciously. Since 2014, finally, the instructions of certain test items have been developed further in the sense that a previously static image has been replaced with a dynamic visualisation of the traffic situation from the driver's perspective. Such dynamic illustrations permit near-realistic presentation of those aspects of a situation (e.g. speeds, action sequences, hazard cues) which are relevant for driving decisions and thus the answers to test questions. This further raises the correspondence between the test demands and the demands of participation in real traffic (see Chapter 4: “Evaluation and further development of the theoretical driving test”).

The traditional test item formats (multiple-choice questions, gap-fill questions) have so far been left unchanged by the outlined optimisation measures – the improvement potential embodied in innovative formats has thus not yet been exploited to the full in the German system of novice driver preparation. In a number of other countries (e.g. Great Britain, the Netherlands), by contrast, various test methods which demand primarily a correct reaction or the correct “driving decision” in the displayed scenario have been developed and implemented over the past decade; at the same time, non-verbal response (e.g. the reaction time before a computer input) is also measured (Genschow, Sturzbecher & Willmes-Lenz, 2013). The importance of such innovative test formats can be seen in the fact that they meet longstanding demands for improvement of the teaching and testing of hazard-related behaviour (Barthelmess, 1976; Hampel, 1977a): A test method which addresses traffic perception and hazard avoidance could in future supplement the “traditional” TDT – with its focus on factual knowledge – and by assessing implicit action knowledge could bridge the content gap to the practical driving test in real traffic (Rüdel, Sturzbecher, Genschow & Weiße, 2011).

On the basis of the above thoughts, it is also possible to formulate the role of the computer in novice driver preparation in more concrete terms, as already indicated in Chapter 1. For the PDT, as an “vehicle-based” test, innovative approaches are already being elaborated to enable the documentation of driving performance and corresponding feedback by way of an electronic test report. With regard to “computer-based” testing, however, it will in future be necessary to distinguish between two components: On the one hand, the traditional multiple-choice format of the TDT can be “upgraded” from the methodical perspective by utilising modern multimedia capabilities (e.g. dynamic situation presentations, random rotation of test items and answer options). On the other hand, the development and testing of innovative test formats (e.g. items permitting evaluations of visual information acquisition and processing speed) will enable competences relating to adequate traffic perception and hazard avoidance to be assessed within the framework of a computer-based “traffic perception test” (hereafter also referred to with “TPT”).

In this chapter, the potential of traffic perception tests with regard to further development of the system of novice driver preparation is to be presented as a basis for analysis and discussion of the essential prerequisites for development and implementation of the corresponding methods. These prerequisites are to be elaborated successively under four key aspects (see Fig. 3), each of which can be assigned independent significance in the context of the overall system of novice driver preparation and thus for the development of appropriate measures: (1) Firstly, it is necessary to determine what is to be understood by "traffic perception and hazard avoidance” from the theoretical and psychological perspectives (“construct description”). (2) Secondly, it must be shown how competence relating to traffic perception and hazard avoidance is to be conveyed in novice driver preparation, including driving school training (“training concept”). (3) On this basis, an assessment concept must be elaborated to describe the test items which are to be used to verify particular elements of the required competence (“test concept”). (4) Finally, it must be explained how a future traffic perception test can be embedded into the system of novice driver preparation (“implementation concept”).

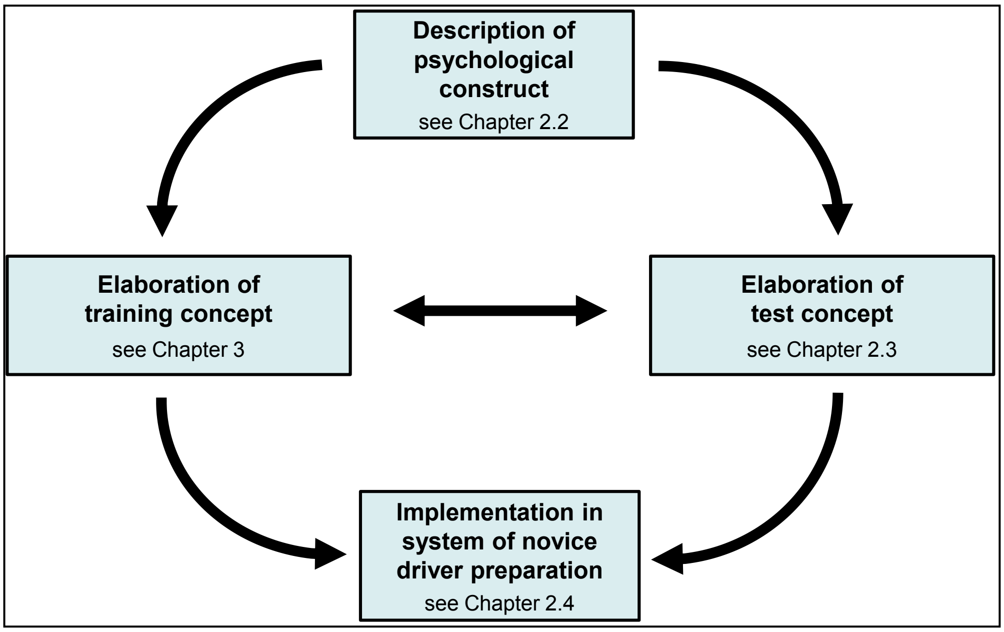



**Fig. 3: Process steps to anchor traffic perception and hazard avoidance in the system of novice driver preparation**

re (1): Research findings indicate that the higher accident risk of novice drivers can be attributed to specific deficits in the acquisition, processing and assessment of action-relevant traffic information (overview in: Horswill & McKenna, 2004). Against this background, Chapter 2.2 discusses the demands placed on drivers with regard to information processing, the current status of research into corresponding novice-specific competence deficits, and models to describe the required competence dimensions.

re (2): The training concepts which appear most promising as a means to improve road safety are those which focus on the computer-assisted transfer and acquisition of competence relating to traffic perception and hazard avoidance in virtual traffic situations, i.e. without exposure to real traffic risks. In Chapter 3, therefore, corresponding training concepts are presented together with empirical findings on the feasibility of training such competence.

re (3): Effective training concepts include provisions for the assessment of learning progress. In the case of novice driver preparation, a distinction must be made between on-going learner assessments during the course of driver training and final assessment in the form of a state-organised driving licence test. As a basis for the elaboration of concepts for continuous learner assessments and a traffic perception test, Chapter 2.3 describes and classifies different approaches for the measurement of traffic perception and hazard avoidance, taking into account empirical findings.

re (4): Didactically meaningful references to the process of driver training, and thus starting points for their embedding into the system of novice driver preparation, must be found for the various forms of driving licence testing – TDT, PDT and in future possibly also TPT. In Chapter 2.4, therefore, the implications for the positioning of different forms of testing within the process of driver training are discussed with reference to international examples.

# 1.6 Theoretical foundations of traffic perception and hazard avoidance

## 1.6.1 Demands and novice-specific competence deficits relating to participation in motorised road traffic

Viewing merely the technical aspects of vehicle handling, the task of driving a motor vehicle does not seem to be an especially demanding proposition. Nevertheless, accident analyses show that novice drivers, in particular, are not seldom overtaxed: "Human error" – as opposed to the influences of technical defects – is by far the most significant cause of accidents (Statistisches Bundesamt, 2014). The reason for this apparent contradiction lies in the fact that, from the perspective of cognitive psychology, the safe driving of a motor vehicle must be seen as a process requiring the continuous solving of "complex problems"[11] (Sturzbecher, 2010). Three demand aspects characterise this problem-solving process:

(1) Firstly, the proper handling of a motor vehicle in road traffic requires the continuous processing of situational information gained from the traffic infrastructure (e.g. road markings, road signs), from other road users and from the driver's own vehicle (e.g. dashboard displays). As the outcome of this information processing, action decisions are derived as a means to master the given traffic demands, which according to Sturzbecher (2010) refer to a "lifeworld domain". Domains are fields of demand profile content in which similar problem solution strategies, knowledge assets and experiences can be applied and for which common normative orientation patterns exist. Following Gruber and Mandl (1996), lifeworld domains can thus be characterised as situations in which the observable phenomena display a high level of complexity and dynamics, as they are subject to an abundance of influencing factors. At the same time, no rules or principles exist with validity for the responses to all possible demand situations. The particular cognitive challenge for the driver, therefore, is the requirement to obtain and evaluate action-relevant information in the lifeworld domain of road traffic: The complex and constantly changing traffic situation generates a diversity of information which must be prioritised and processed further.

(2) A second challenge for drivers derives from the fact that various cognitive action processes (e.g. traffic observation in search of information, navigation) must be realised simultaneously with motor actions (e.g. braking or gear-changing). The mastering of such multiple demands requires the execution of different primary driving actions, possibly even in conjunction with additional secondary activities (e.g. conversation with a passenger), for which it is necessary to develop corresponding action routines in the sense of driving experience. As long as no such routines are available to relieve the driver's cognitive working capacities, even simple dual tasks can lead to a reduced quality of performance (Kahneman, 1970; Koch, 2008); after adequate practice, however, the level of performance displayed when processing dual tasks is equivalent to that achieved for each single task (Spelke, Hirst & Neisser, 1976).

[11] According to Funke (2003), so-called "complex problems" are characterised by the following five system attributes: Complexity (i.e. a large number of situation variables must be taken into account); connectivity (i.e. consideration must be given to the mutual dependencies between framework conditions); dynamics (i.e. the situation continues to develop further even without action on the part of the person concerned); intransparency (i.e. the information required for problem-solving decisions is not accessible in its entirety, for example for time reasons) and polytely (i.e. it is also necessary to weigh up conflicting variables).

(3) A third challenge when driving a motor vehicle can be seen in the fact that information processing and the subsequent determination and implementation of an action decision must be accomplished within more or less narrow time limits. The driver is at any moment an inherent element of the complex and dynamic traffic situation, and is thus forced to display behaviour of one kind or another (Leutner, Brünken & Willmes-Lenz, 2009). Consideration of the fundamental physical correlations already illustrates how quickly the time window for appropriate action in a certain situation opens and closes. A vehicle travelling through a residential area at the relatively slow speed of 30 km/h, for example, already covers a distance of around nine metres in just one second. This period of one second corresponds approximately to the assumed reaction time, and it thus follows that, after the driver identifies a hazard-relevant information cue, the vehicle will travel a further nine metres at its present speed before the brakes are actually applied. This example underlines the major significance of the factor “time” for safe driving, and also indicates how acceleration of a driver's reaction by just fractions of a second can already yield considerable safety gains.

From the demands outlined above, namely processing of a flood of information and coordination of different forms of action under time pressure, it becomes clear that safe driving in road traffic represents an enormous cognitive and psychomotor achievement. Thorough preparation is imperative to enable novice drivers to manage such demands independently, and they must themselves support the success of this preparation by building up relevant driving experience over a longer period: It is known from various empirical studies that novice drivers face the highest accident risk of their whole driving career when they first start solo driving. Subsequently, the accident risk sinks rapidly with increasing driving practice (Maycock, Lockwood et al., 1991; Schade, 2001). A number of aspects which could explain the safety-enhancing potential of increased practical driving experience were identified by Deery (1999) in an analysis of research findings on novice-specific driving competence deficits and on differences in driving performance between novice and experienced drivers (see Fig. 4).

**Novice drivers …**

… perceive situations less holistically (piecemeal and independently of the context).
… perceive hazards less quickly.
… focus on a narrower area when performing horizontal scans.
… concentrate their attention on a shorter stretch of the road in front of their vehicle.
… check mirrors infrequently.
… use their peripheral vision inefficiently.
… fix their gaze on fewer objects.
… fix their gaze more on stationary objects.
… are more likely not to perceive hazards altogether.
… are more likely to underestimate risks.
… consider themselves to be more skilful than an “average driver”.
… are overconfident with regard to successful application of their driving abilities.
… judge their individual risk to be lower than that of other young drivers.
… display greater risk acceptance.
… learn vehicle handling skills quickly, but require more time to develop higher-order cognitive skills.

**Fig. 4: Novice-specific driving competence deficits (based on Deery, 1999)**

It can be taken from Fig. 4 that, at the beginning of their driving careers, novice drivers lack above all adequately developed skills relating to visual information acquisition and processing. In this context, Horswill and McKenna (2004) point out that the ability to recognise potential risks in road traffic (“hazard perception“; see Chapter 2.2.3) is a central element of driving competence and correlates closely with accident risk. When seeking to improve novice driver safety, therefore, it seems expedient to address the cognitive process sequences of traffic-related information acquisition and evaluation by way of correspondingly targeted training and testing approaches during the preparation for solo driving (see Chapter 3 and Chapter 2.3). The following sections present a number of cognitive psychology notions and concepts which can be used to describe acquisition of the necessary cognitive prerequisites for the driving of a motor vehicle more precisely; they serve furthermore as a basis for the further discussions in this report.

*Visual information reception*
When driving a motor vehicle, visual information reception – as opposed to other forms of perception such as auditive perception or the sense of balance – represents the principal sensory channel for the acquisition of information (Sivak, 1996; Hole, 2007). The visual system used by the human eye to gather information is characterised by a binocular field of vision covering a range of almost 180 degrees horizontally and 130 degrees vertically. The section of this field of vision in which objects can be observed at the highest resolution and in maximum detail (the area of so-called "foveal vision"), however, covers merely 1 to 2 degrees (Vollrath & Krems, 2011). Visual scanning is performed as a sequence of fixations and saccades, i.e. moments during which our gaze is held at a particular location to enable the analysis of visual stimuli, and intermediate periods during which this focus jumps from one location to another. Between one and three fixations occur per second (Goldstein, 1997).

*Perception processes*
Zimbardo (1992) distinguishes three stages in his description of (visual) perception processes (see Fig. 5): At the initial "sensory stage", a physical object present in the surrounding environment ("distal stimulus") is projected as an image onto the retina ("proximal stimulus") and produces a corresponding sensory sensation. At the second stage of "perceptual organisation", sensations are transformed into a set of patterns and forms to enable "perception" or, in other words, internal representation of the external stimuli (e.g. two parallel vertical lines joined at their centre with a third horizontal line are recognised as the letter "H"). The third "classification" stage, finally, then assigns the identified image to a known category.

Such (visual) perception processes can be triggered in different manners. A distinction is made between "bottom-up" (or "data-driven") and "top-down" (or "conceptually driven") processes, though both are in constant interaction with each other (Zimbardo, 1992). In the case of a "bottom-up" process, perception is triggered by physical or sensory stimuli received from the environment; attention could be attracted, for example, by conspicuous properties of an object such as its colour or movement. In a "top-down" process, by contrast, the perception processes are already initiated by existing memory content, e.g. knowledge of the hazardous nature of a particular situation and the expectation that it will occur.

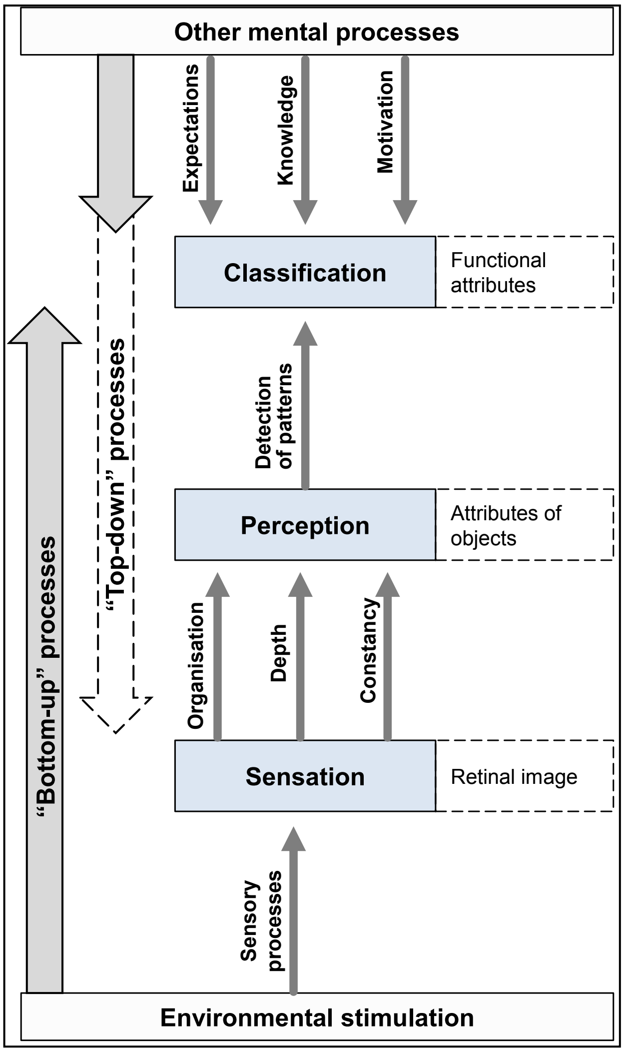



**Fig. 5: "Top-down" and "bottom-up" perception processes (based on Zimbardo, 1992)**

*Cognitive schemata*

The terms "cognitive schema" and "script" are used in cognitive psychology to describe forms of knowledge organisation (Bartlett, 1932; Schank & Abelson, 1977). While the notion of cognitive schemata generally refers to the organisation and storing of explicit knowledge, the script concept is instead associated with prior knowledge of situational scenarios and with situation-related expectations. In this respect, scripts also represent mental knowledge structures or schemata, but refer instead to procedural knowledge and action expectations in the sense of a storyboard. Under this approach, scripts can be taken to represent the storing of our everyday knowledge of events and their significance.[12] The script contents describe the roles of the agents involved, the initial and final states of the event, contributing facts or details, and the sequence of actions (Kluwe & Spada, 1981). Alongside these concrete components, scripts – as classes of stereotypical social events – include also empty "slots" to accommodate special situational circumstances (Silbereisen, 1987). Scripts are activated by a specific situation, adapted to the current situational context by filling the gaps left by the slots, and then serve as instruments of experience-driven action control (Kluwe & Spada, 1981; Nelson, 1986). The advantage of cognitive schemata and scripts with regard to efficient information processing lies in the possibility

[12] "A 'script' is an elaborate causal chain which provides world knowledge about an often experienced situation ..., scripts are predetermined sequences of actions that define a situation" (Schank, 1975, p. 264).

to incorporate new information; in this way, through combination with stored prior knowledge and prior assumptions, greater meaning can be drawn even from incomplete information (Zimbardo, 1992).

*Situation awareness*
A further approach to describe memory content is the concept of "situation awareness". This approach refers explicitly to dynamically changing systems and is used primarily in research addressing human-machine interaction. Endsley (1995) defines situation awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (p. 36). In accordance with this definition, three levels of situation awareness are distinguished: (1) "Perception", (2) "Comprehension" and (3) "Projection" (see Fig. 6).

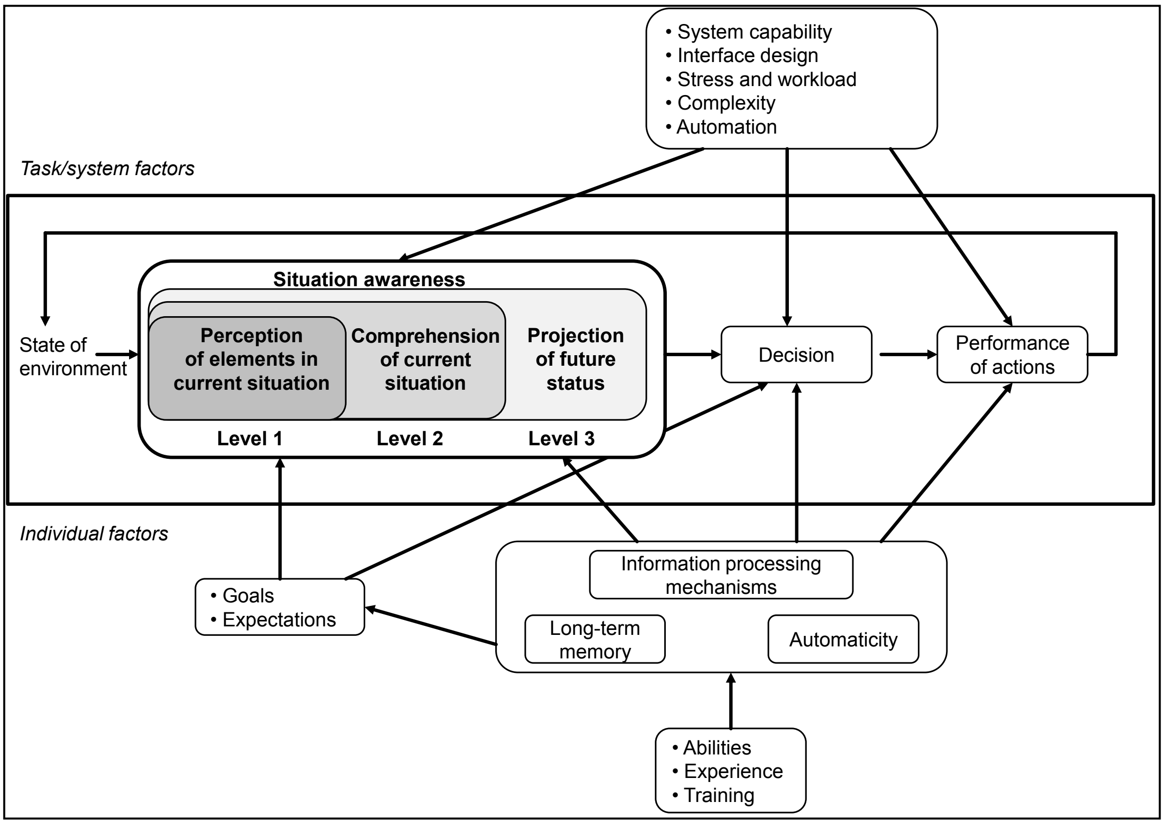



**Fig. 6: Model of situation awareness (Endsley, 1995)**

On the first level, the status, attributes and dynamics of relevant elements in the environment are perceived. A pilot, for example, needs to know the positions, status and course of other aircraft, as well as the status and course of his own aircraft. On the second level, the perceived elements (and their attributes) from the first level are combined into a holistic image of the environment, whereby the significance of particular objects and events is identified. On the third and highest level of situation awareness, finally, future actions of the elements in the environment are projected. This is achieved on the basis of knowledge of the status and dynamics of the individual elements, together with an understanding of the situation as a whole.

The three described levels of situation awareness build upon each other; mastering of the first and second levels is in each case a prerequisite for progression to the subsequent level (Wickens, 2008). Endsley (1995) emphasises that the construct "situation awareness" is to be kept clearly separate from the immediate subsequent decisions and actions: It is true that well-developed situation awareness is a necessary condition for successful decision-making and for actions appropriate to the demands to be met, but it is not already sufficient in itself (Wickens, 2008). Accordingly, even a person with pronounced situation awareness can still make wrong decisions or otherwise act dysfunctionally, i.e. display poor performance. Furthermore, situation awareness must be distinguished from knowledge stored in long-term memory. Situation awareness describes merely the state of a person's knowledge at a particular moment in time. This state of knowledge is, in certain respects, “by nature” temporary and time-dependent; it exists only for a few seconds or minutes. Knowledge contents (schemata, scripts) stored in long-term memory, on the other hand, remain valid for days or years (Wickens, 2008).

In the following Chapter 2.2.2, it is now to be discussed how the described psychological concepts relating to information acquisition and processing can be applied to the context of participation in motorised road traffic, and how they can be utilised for novice driver preparation.

### 1.6.2 Models of traffic perception and hazard avoidance: Approaches from the field of traffic psychology

The demands to be met by a description of driving and traffic competence can be derived above all from the competence models used in school education research. In this connection, reference must be made to the distinction between competence structure models and competence level models, as presented by Klieme and Leutner (2006): Competence structure models reflect the various dimensions of competence which can be differentiated in a specific field, whereas competence level models establish correlations between concrete situational demands and the scope of competence required to master those demands. These models are nevertheless not mutually exclusive, and are ideally even complementary (Koeppen, Hartig, Klieme & Leutner, 2008). Sturzbecher and Weiße (2011) consider it desirable to use both models to describe driving competence, as this would incorporate both the various content-based dimensions of driving competence (partial competences) and the possible levels of attainment of such partial competences among novice drivers. At the same time, however, the authors point out that it is scarcely possible to develop a complete or general model of driving competence, because – in accordance with the purpose served – competence models always focus on a very specific aspect of driving competence.

In the present innovation report, the aspect of traffic perception and hazard avoidance is placed in the foreground. At this point, therefore, safety-relevant demands and the required competences corresponding to this aspect (“demand analysis”) are to be drawn from the current literature relating to information processing and action regulation in connection with the driving of a motor vehicle. The expected outcome is a learning-objective-referenced definition of the construct “traffic perception and hazard avoidance”, along with descriptions of its relevant components, as a basis for the development of corresponding training modules and a traffic perception test.

The literature research revealed four task analysis models which address demands and competences in the field of traffic perception and hazard avoidance. These models, which were each developed with different primary objectives – are to be described in more detail in the following sections, together with findings relating to their validity.

*“Model of information processing in traffic situations” (Sturzbecher, Kammler & Bönninger, 2005)*

The authors of this model took up the “model of social information processing”[13] which had been presented by Crick and Dodge (1994) and applied it to information processing mechanisms in the (social) context of motorised road traffic. The model elaborated in this way is understood as a basis from which to derive learning content and corresponding teaching strategies for driving school training, as well as test contents and testing strategies for the subsequent driving licence tests, from the perspectives of education and test psychology. It describes six action stages as expression of the social-cognitive demands to be met by a driver in terms of behaviour regulation in a given traffic situation (see Fig. 7).

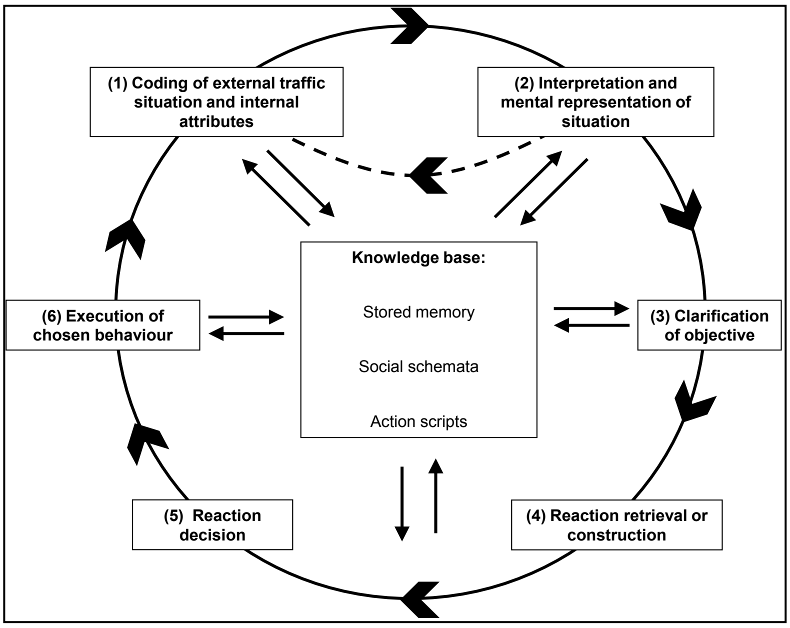



**Fig. 7: Model of information processing in traffic situations (based on Crick & Dodge, 1994)**

The following action demands were distinguished by Sturzbecher et al. (2005) and explained by way of the necessary action steps:

1. Coding of the external traffic situation and internal attributes (through the perception of (social) information, the focusing of attention and targeted information searches under recourse to existing schemata stored in long-term memory);
2. Interpretation and mental representation of the situation (through analysis of the current situation and conclusions drawn from comparisons with previously experienced situations);
3. Clarification of the objective (decision between further development of the original objective and selection of a new objective to achieve the desired action result);
4. Reaction retrieval or construction (retrieval of known reaction alternatives from memory or else construction of a new reaction possibility);

[13] This model describes human behaviour in social situations from a social-cognitive perspective.

5. Reaction decision (evaluation of the reaction alternatives on the basis of the expected consequences, competence expectations and the appropriateness of the reaction concerned; determination of the optimum reaction), and
6. Execution of the chosen behaviour (selection of an appropriate option from the available behaviour repertoire and proper realisation of this behaviour).

Sturzbecher et al. (2005) saw the described model above all as a theoretically founded framework for the designing of application- and action-referenced test items for the theoretical driving test: The model offered a possibility to match test demands to the perception, thought and decision processes which are actually applicable when driving in road traffic. There was no empirical validation of the adapted model, however.

*"Model of hazard cognition" (Schlag, 2009)*
Taking up the references to social-cognitive theory which Sturzbecher et al. (2005) described in connection with information processing in traffic situations, it can be deduced that a driver must constantly strive to produce an adequate internal representation of the external environment. The task is essentially to develop a "picture" of the traffic situation – of which the driver is himself an inseparable element; given the dynamic nature of road traffic, this picture is furthermore subject to permanent modification and correction.

Schlag (2009) also describes constant information acquisition and processing within the framework of (traffic) perception processes as a basis for (driving) behaviour decisions and appropriate actions (in road traffic). The psychologically effective action conditions are here not the actual conditions of the external environment, but rather their cognitive representation. This, in turn, is dependent not only on the perceived external stimuli ("bottom-up"), but also on personal factors such as motivation, experience or expectations ("top-down") (see Chapter 2.2.1). In his "model of hazard cognition" (see Fig. 8), Schlag (2009) describes information processing with regard to traffic-related hazards as a six-stage process.

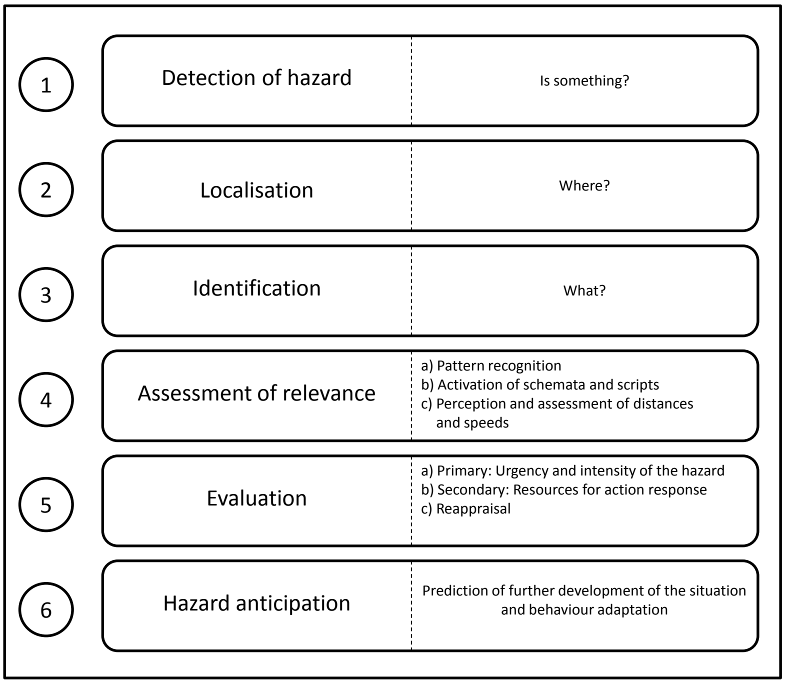



**Fig. 8: Model of hazard cognition (based on Schlag, 2009)**

The first three stages of the model described by Schlag (2009) refer to the "detection", "localisation" and "identification" of a hazard. Schlag apparently views the initial "detection" stage as a "bottom-up" process step, which represents a "counterpole to the perception of resources (e.g. self-assigned driving abilities) and motives for their utilisation" (ibid., p. 53); objects can be detected more easily if they stand out from other stimuli. The importance of the second stage of "localisation" is seen above all in connection with other road users, who can still be perceived (whether visually or acoustically) even if they approach from behind.[14] The third stage of "identification", according to Schlag, concerns determination of the manner of the hazard and what it could mean for the driver's own behaviour.

The identification stage (stage 3) is followed by stages 4 "Assessment of the relevance" and 5 "Evaluation". For Schlag (2009), these stages refer to the driver's understanding and evaluation of both the hazard situation and his own action possibilities ("How do I judge the situation? Do I have appropriate schemata and scripts, i.e. action-referenced mental models? Can I transfer past experience?").[15] For the evaluation stage, the driver's assessment of the hazard in relation to his possible action response is also important: An objectively dangerous situation, for example, could be viewed as a challenge rather than a hazard to be avoided if the driver is convinced of his ability to handle the situation. At the sixth stage of "hazard anticipation", finally, the further development of the situation must be forecast, taking into account the driver's own action tendencies and those of the other road users. This reference to other road users places particular demands on the ability to assume other perspectives, which, alongside the timely and correct identification of hazard-relevant information cues, is especially important for successful hazard avoidance.

The described model was already applied in connection with a presentation of empirical findings on driving behaviour in different light and visibility conditions (Schlag, Petermann, Weller & Schulze, 2009) and in an analysis of the acoustic perception of vehicles with alternative drive systems by vulnerable road users (Hagen, Schulze & Schlag, 2012). There has been no explicit reference to either driving licence testing or the process of learning to drive, and no findings exist to date on the validity of the model.

*"Model of the processes underlying driving behaviour in response to potential hazards in road traffic"(Deery, 1999)*

Deery based his model on the published results of studies addressing novice-specific competence deficits. The model elaborated in this way (see Fig. 9) is intended to depict the perceptual and cognitive processes surrounding driving behaviour in the context of traffic hazards. The aim is not to provide a complete description of the processes which affect decision-making and risk-taking when driving a motor vehicle, but rather to determine the main elements which could influence a driver's behaviour with regard to hazard recognition and avoidance. Deery treats his model as a framework within which the relevant literature and research findings can be categorised and subjected to critical appraisal,

[14] In models of visual information reception – and thus also in the case of Schlag's model – two qualitative stages must be distinguished: The objects in a given space are first localised at a pre-attentive stage, and only then identified within the framework of a subsequent attentive stage. In an empirical study of the eye movement of users of a PC-based driving simulator, Velichkovsky et al. (2002) showed that the manner and duration of fixations which led to recognition of an immediate hazard differed from other gaze patterns.

[15] Following the "stress model" elaborated by Lazarus (1966), information is analysed in three steps at the "evaluation" stage. According to Lazarus, stress arises from factors in the environment ("stressors"), which, in the course of a "primary assessment", may be classified as positive, irrelevant or (potentially) dangerous. In the latter case, the resources available to deal with the situation are analysed, in preparation of a secondary assessment. If the resources are found to be inadequate, this triggers stress, which must then be overcome by way of suitable strategies ("coping").

and thereby serve to identify opportunities for further research in respect of the individual model components.[16]

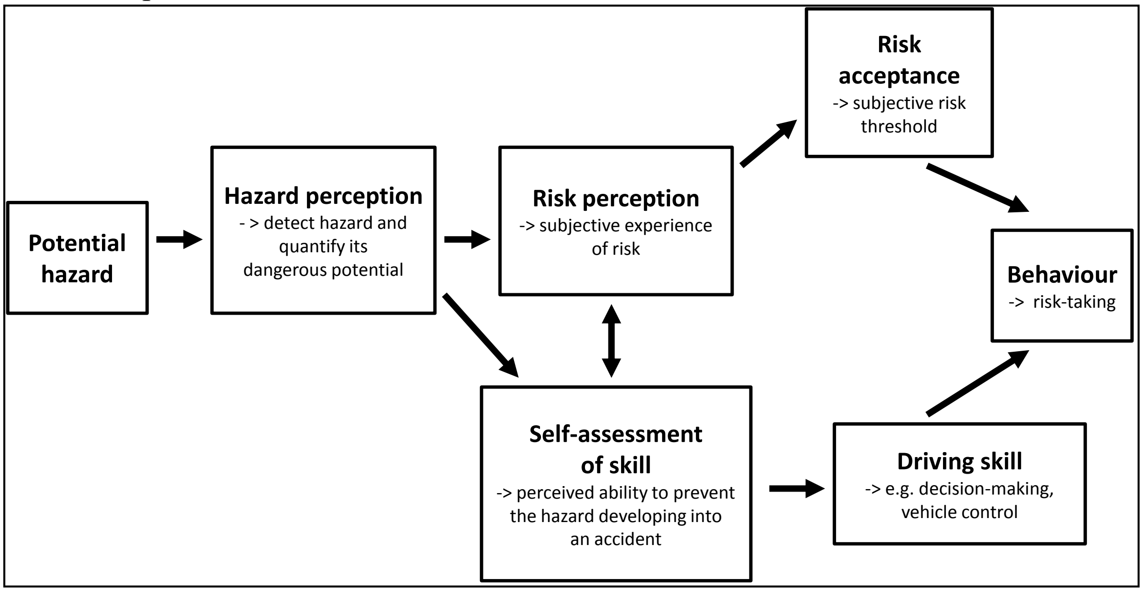



**Fig. 9: Model of the processes underlying driving behaviour in response to potential hazards in road traffic (Deery, 1999)**

Deery's (1999) model takes into account the fact that the concepts of “risk” and “hazard perception” also reflect subjective experiences on the part of the driver, and that these experiences must be distinguished from objective risk. Referring to Brown and Groeger (1988), he describes “objective risk” as a quantity which can be estimated, wherein the reliability of the estimate is dependent on the information which is (or can be) used in the calculation. “Risk perception”, by contrast, is governed by subjective experience. It is dependent on both the available information regarding potential hazards in the traffic environment, and information on the ability of the driver (and the technical capabilities of the vehicle) to avoid those potential hazards.

Taking up the distinction between subjective and objective risk, Deery (1999) describes the model component “hazard perception” as the process of identifying a hazard and quantifying its “objective” risk. This component is thus assigned elements of competence relating to information acquisition, and here in particular those connected with gaze pat-

[16] To eliminate possible ambiguities arising from the often inconsistent use of the terms “hazard” and “risk” in research literature and everyday language, the present report adopts and uses the following concept definitions: The term “hazard” is deemed to provide a qualitative description of (dangerous) traffic situations (i.e. objects or object constellations) or events which hold a certain injurious potential on account of their specific aspects (properties, features, states). This understanding is in line with the WHO definition of “hazard” as an “… inherent property of an agent or situation having the potential to cause adverse effects when an organism, system, or (sub)population is exposed to that agent” (IPCS, 2004). The term “risk”, which the WHO describes as the “… probability of an adverse effect in an organism, system, or (sub)population caused under specified circumstances by exposure to an agent” (IPCS, 2004), on the other hand, refers to the quantifiable probability of the occurrence of an actual accident. In this connection, a distinction is to be made between an “objective risk”, which can be determined by applying scientific methods (e.g. expert surveys, rating methods, accident analyses), and a “subjective risk”, which embraces individual experience- and behaviour-related aspects of risks, including the weighing up of possible negative consequences against the expected gain or benefit (e.g. risky behaviour, so-called “risk-taking”). With regard to a possible accident, therefore, the two probabilities of occurrence – in the sense of an objectivised risk on the one hand and a subjectively “coloured” assessment of risk on the other – can thus differ fundamentally.

terns (e.g. number of objects pinpointed by way of fixation, search strategies). The component "risk perception", on the other hand, concerns the driver's assessment of the existing risk in a given situation on the basis of available information. A further component which is mentioned is "self-assessment of skill". This refers to subjectively perceived abilities with regard to hazard avoidance, i.e. the driver's own capacity to prevent a hazardous situation developing into an accident. The component "risk acceptance" then describes the individual risk threshold which the driver is willing to accept; this is expressed in the extent to which degrees of freedom in driving behaviour (e.g. speed selection, route selection) are utilised ("self-paced task"). These components thus belong to the motivational and other personality-related aspects of driving (e.g. impulsiveness, "sensation seeking") which determine the individual risk threshold and the choice between available action alternatives. The component "driving skill", finally, refers to actual control of the vehicle.

Evidence for the described model components is to be found in current research results and permits the construction of various hypotheses on the causes of risky driving behaviour.[17] The validity of the model has not been investigated.

*"Model of the process of response to hazards" (Grayson, Maycock, Groeger, Hammond & Field, 2003)*
The authors developed a four-stage process model to study the interactions between possible influencing factors as the foundation for an explanation of novice driver risk. The postulated stages serve to reflect the performance differences between individual drivers in their reactions to hazards (see Fig. 10). The first stage is "hazard detection"; if this is not accomplished successfully, the risk that the existing hazard will lead to damage or injury is increased. Once the driver has detected a hazard, he must assess the threat posed by that hazard ("threat appraisal"). If it is deemed necessary to react, a suitable reaction must be selected in accordance with the situational circumstances ("action selection"). To avoid the hazard, finally, the chosen action must actually be performed ("implementation") – whether or not this implementation is successful will then depend on the abilities of the driver concerned.

[17] Deery (1999) illustrates this with the example of a driver who continues to travel at a relatively high speed despite the fact that a pedestrian is crossing the road ahead: Firstly, the driver's risk perception may be poor (e.g. he fails to judge the distance to the pedestrian or the braking distance of the vehicle correctly). Secondly, he may overestimate his ability to deal with the situation. Thirdly, he may possess relatively poor levels of driving skill and thus take excessive time to detect the presence of the pedestrian, make the decision to slow down and then actually apply the vehicle brakes. Finally, he may possess a high risk threshold, i.e. he decides not to brake and instead accepts the risk of an accident.

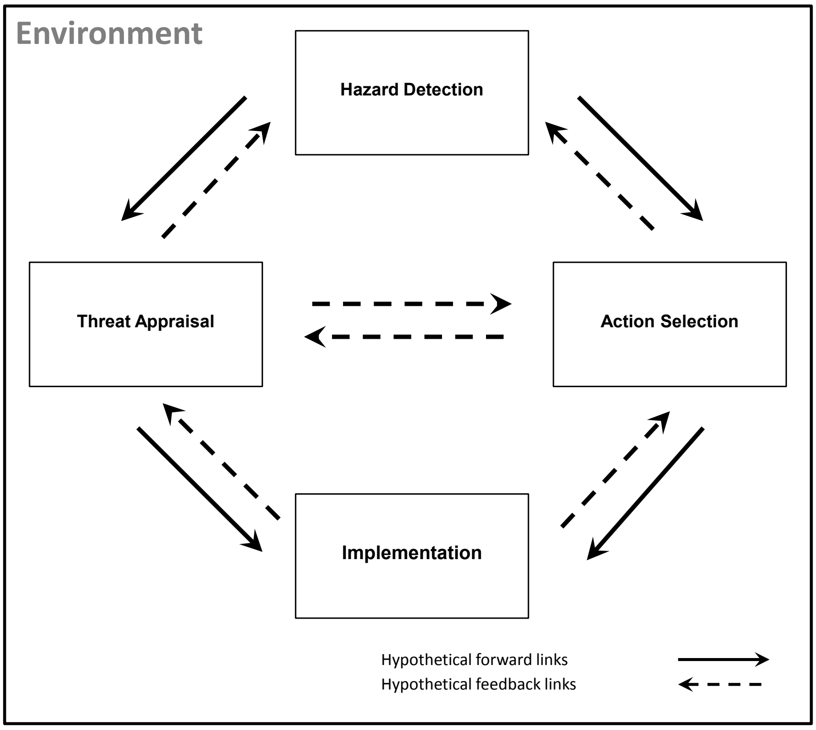



**Fig. 10: Model of the response to hazards (Grayson et al., 2003)**

Grayson et al. (2003) presented a number of hypotheses regarding the interactions of the individual model components and tested these hypotheses empirically.[18] Correlations between the model components and self-reported accident involvement were revealed in various variables. It was found, for example, that – compared to drivers with prior accident involvement – persons with an accident-free driving record

- achieved higher questionnaire scores for the factor "thoroughness" in the context of "hazard detection",
- gave a lower self-assessment of their vehicle handling skills in hypothetical traffic situations and showed reduced risk acceptance and impulsiveness in the context of "threat appraisal",
- displayed faster reactions in finger-based reaction tasks representing the stage of "implementation" (this remained true when controlling for the factor "age").

[18] To this end, a computer-based test battery ("Computerised Assessment of Driving Skills – CADS") was developed with appropriate methods to assess a variety of driving performance and personality variables. For the component "hazard detection", the test battery comprised items requiring the assessment of traffic scenarios, as well as the NFER hazard perception test (Grayson & Sexton, 2002). "Threat appraisal" was operationalised, for example, by way of a questionnaire on self-assessed driving ability, scales relating to "impulsiveness", "readiness to take risks" and "experience of stress" (from the "Differential Personality Questionnaire" by Tellegan, 1992), as well as a scale relating to internal and external control attitudes. With regard to "action selection", aspects of the reaction time to given stimuli (e.g. simple reactions involving a manual action or use of a pedal; timed selections, where the appropriate reactions were to be selected from a set of defined options), as well as aspects of intelligence and selective attention and inhibition, were measured. The items serving to test "implementation" were above all motor process tasks (e.g. realisation of a requested steering action using the steering wheel in front of a monitor; "tracking" of a moving target with a cursor controlled in two-dimensional space by way of a pedal and a steering wheel). In addition to the diversity of variables representing the individual components of the model, the study participants (N=404) also provided information on past accident involvement.

The next step in the study placed the test performances recorded under laboratory conditions in relationship to the driving skills displayed in real traffic situations. To this end, the driving behaviour of a sub-sample of the study participants (N=100) was observed during a test drive in their own vehicles and recorded on video for subsequent evaluation. Each participant completed a 16-kilometre test drive along a standardised driving route. At selected points characterised by particular demand situations (e.g. bend with limited visibility, different types of road), the drivers were furthermore asked to judge the perceived hazardousness and difficulty of the traffic situation. Their driving speed was also measured.

The study results show significant correlations between the performance shown by drivers in the laboratory tests and the observer judgements of their actual driving behaviour. Better perception in the hazard perception test, for example, could be associated with relatively low driving speeds in real traffic and better observer assessments in respect of foresighted driving. A relatively high driving speed could be linked to a negative observer assessment – the driver was described as displaying an unsafe manner of driving and inadequate attention, for example – and a tendentially higher level of reported accident involvement. With regard to the driver judgements of real traffic situations (hazardousness, difficulty), however, the presumed correlations to performance in the laboratory tests were for the most part not confirmed. Possible reasons were seen in methodical differences between the measuring procedures and uncontrolled confounding variables in the real traffic situation.

To corroborate the validity of their model, Grayson et al. (2003) conducted a further study in which they surveyed a larger sample of drivers (N=1340) by way of a questionnaire addressing selected variables of the CADS test battery and the respondents' past involvement in accidents. The revealed correlations between the survey variables and reported accident involvement are overall indicators for the good criterion validity of the model components (e.g. participants who judged traffic scenarios to be more hazardous were subject to a lower accident risk).

Viewed as a whole, the model elaborated by Grayson et al. (2003) is substantiated by the empirical studies; it thus offers a suitable framework for the derivation of relevant structural components of traffic perception. It must be noted critically, however, that the empirical methods assigned as variables to the individual model components are – it would seem – not always suitable to obtain a conclusive, component-specific statement. The hazard perception test and the items requiring the assessment of traffic scenarios, for example, are assigned to the model component "hazard detection", but also incorporate certain demands relating to "threat appraisal".

### 1.6.3 Description of the construct "traffic perception and hazard avoidance"

In the following, the demands to be placed on novice drivers in respect of traffic perception and hazard avoidance are to be elaborated. The objective of the demand analysis is to determine possibilities to convey and measure competence within the framework of driving school training and a subsequent driving licence test, and to categorise these demands in an overarching theoretical construct. To this end, the models of traffic perception and hazard avoidance skills presented in Chapter 2.2.2 are first subjected to comparative analysis (see Table 1).

It can be taken from Table 1 that, notwithstanding the different objectives pursued when developing the models, the components of the aforementioned models display a certain fundamental correspondence in terms of content. The extent to which the model components are differentiated, however, varies strongly between the individual authors. The element of “assessment”, for example, is considered an important component in all models. In the model presented by Deery (1999), however, this aspect is depicted in particular detail, in that distinctions are made between the subjectively perceived risk (“risk perception”), the driver's self-assessment of his skills to act accordingly and the risk which the driver is prepared to accept. Grayson et al. (2003), on the other hand, summarise these aspects under a single model component. The models elaborated by Schlag (2009) and Sturzbecher et al. (2005) display comparatively greater differentiation with regard to the (visual) recognition and structuring of situation attributes. Both models point to the internal representation of situation attributes as an aspect of particular significance. In the models by Grayson et al. (2003) and Sturzbecher et al. (2005), furthermore, the process of active avoidance of a hazard by way of corresponding behaviour is subdivided into selection of an appropriate action and actual realisation of the chosen option.

On the basis of the model-specific differentiations described above, a series of structural components can be derived for the construct “traffic perception and hazard avoidance”; these components are entered in the left-hand column of Table 1 as demands to be met by drivers. The individual demands refer to “observation”, “localisation” and “identification” of the situation, the driver's “assessment of the hazard”, “assessment of own abilities” and “weighing up of risk”, and finally the ensuing “decision” and “action”.

**Tab. 1: Model analysis and determination of demands to be mastered by novice drivers**

| | | **Overview of model components relating to traffic perception and hazard avoidance** | | | |
|---|---|---|---|---|---|
| | | Sturzbecher et al. (2005) | Schlag (2009) | Deery (1999) | Grayson et al. (2003) |
| **Demands placed on drivers** | **Observation** | Coding of external traffic situation and internal attributes | Detection (is something?) | Hazard perception | Hazard detection |
| | **Localisation** | | Localisation (where?) | | |
| | **Identification** | Interpretation and mental representation of situation | Identification (what?) | | |
| | **Assessment of the hazard** | Clarification of objective | Assessment of relevance | Risk perception | Threat appraisal |
| | **Assessment of own abilities** | | | Self-assessment of skill | |
| | **Weighing up of risk** | Reaction retrieval or construction | Evaluation | Risk acceptance | |
| | **Decision** | Reaction decision | Anticipation | Driving skill | Action selection |
| | **Action** | Execution of chosen behaviour | | | Implementation |

In the following, the aforementioned demands relating to traffic perception and hazard avoidance are to be described in more detail, wherein corresponding reference is made to the model descriptions of the individual authors. At the same time, consideration is given to further thoughts which appear relevant for descriptions of the components with a view to development of a training concept (see Chapter 3) and a diagnostic approach (see Chapter 2.3).

*Demand component “Observation”*

At the “observation” stage, the constantly applicable demand is to acquire information on a specific traffic situation – including any situational cues to potential hazards. In line with the discussions presented by Zimbardo (1992) and Schlag (2009), this can be realised in the form of a "bottom-up" or "data-driven" process, in which hazard cues are perceived because they capture the attention of the driver quasi-automatically in the context of a complex traffic situation as salient objects (e.g. due to their colour, size, direction of motion, etc.). Alternatively, the driver could also search consciously for cues to potential hazards ("top-down" or "conceptually driven" process), in which case the orientation of the search is determined by the driver's prior knowledge (e.g. schemata and facts anchored in his memory). A near-accident with an inattentive pedestrian, for example, may lead to more precise observation of pedestrians in future situations.

Observation of the traffic environment, as a prerequisite for the recognition of relevant situation attributes, represents the first step in traffic perception and hazard avoidance: In motorised road traffic, a multitude of stimuli compete for the driver's limited cognitive resources at any given moment. It is thus of elementary importance to (visually) detect potential hazard cues through observation of the traffic situation, and then to be able to process and categorise the pertinent data. The observation demand thus covers the whole process of acquiring visual information. At the same time, however, the limited availability of cognitive and time resources calls for efficient observation strategies ("scanning" of the environment, use of mirrors), so as to allocate the available resources above all to those situation attributes with a high information content in the context of hazard avoidance. Following the explanations of Deery (1999), novice drivers here display certain competence deficits compared to their more experienced counterparts (see Chapter 2.2.1).

*Demand component "Localisation"*
The purpose of localisation is to assign objects from the traffic environment to a particular position in space. Given the dynamic nature of road traffic, this is a continuous process: Any localised position in space is immediately subject to change, firstly due to the dynamics of other road users, and secondly as a result of the observer's own motion. The positions and orientations of the objects in space must thus be determined repeatedly, at the same time taking into account the dynamic changes in the driver's own situation relative to all other objects. This demand is also reflected in the first level of situation awareness (Endsley, 1995; see Chapter 2.2.1). The entire 360-degree range around the vehicle is significant. This range must be anchored in a "mental model" of the traffic situation, i.e. represented internally and constantly refreshed (Sturzbecher et al., 2005). The demands of localisation can be illustrated by considering the example of mirror use: Although an object recognised in the mirror lies within the driver's field of vision (i.e. apparently in front of the driver), it is in fact positioned either behind or alongside his own vehicle. The driver is thus required to transfer the information acquired from the mirror into a context reflecting the real circumstances. Localisation demands also become evident in connection with the so-called "blind spot", where an object temporarily disappears from the driver's field of vision, but physically still constitutes a real hazard. The discrepancy between the visible extract of reality and the broader entirety of the traffic environment can only be bridged with a mental model of the traffic situation.

*Demand component "Identification"*
The identification of objects requires that they be compared against memorised representations in respect of their specific attributes ("top-down"-driven) and assigned to available categories. This is realised with varying levels of precision and is dependent on the prior knowledge and expectations of the driver. The objects of a particular category, for example "pedestrians", can be qualified with a diversity of possible attributes. The more precisely the attributes of a pedestrian (e.g. age, gender, running or walking, etc.) can be determined in a particular situation, the more precise the assumptions derived with regard to that pedestrian's possible behaviour and its relevance for one's own necessary reaction

(e.g. the identification of a walking adult leads to different expectations with regard to further development of the situation than would be the case for a running child). The importance of identification is also underlined by the example of motorcyclists (Kühn, 2008): Accident analyses show that motorcycle accidents in which a car is also involved occur frequently in situations in which the car is turning left and a motorcyclist is either travelling in the opposite direction or else overtaking, i.e. the car driver has violated the motorcyclist's right of way (Noordzij, Forke, Brendicke & Chinn, 2001). These accidents can be attributed in part to the fact that, although the motorcyclist was seen, his speed and distance were not judged appropriately (Clarke, Ward, Bartle & Truman, 2004). There are thus apparently deficits in the available knowledge base with regard to the characteristics of motorcyclists as participants in road traffic. Only an elaborated knowledge base − organised in the form of cognitive schemata and scripts (see Chapter 2.2.1) – enables the diversity of road traffic information to be classified and perceived objects to be assigned attributes which permit precise behaviour forecasts.

*Demand component "Assessment of the hazard"*
Assessment of the situation in hand is a further essential demand relating to traffic perception and hazard avoidance. Schlag (2009) speaks in this context of "primary assessment" of the traffic situation with regard to the urgency and intensity of the potential hazard. According to Deery (1999), the process of hazard assessment relies on elements of "objective" experience, which is applied in appropriate manner to the information drawn from the traffic situation. If, while driving on the motorway, for example, the brake lights of the preceding vehicles are seen to come on and a tailback ahead is thus identified as a significant traffic situation, this situation should usually be assessed as very urgent due to the high speeds involved. Inappropriate assessments of hazard potential from the information available in the traffic environment can be deemed a novice-typical deficit and are expressed, for example, in the fact that greater hazard potential is associated with stationary objects than with moving objects (see Chapter 2.2.1).

*Demand component "Assessment of own abilities"*
This demand component refers to a driver's assessment of the abilities and resources at his disposal to avoid a hazard. The higher the assessment of his own ability to master a situation, the lower the perceived level of risk emanating from the situation concerned. The assessment of different action alternatives similarly influences the driver's decision in favour of a particular option. Schlag (2009) describes this assessment step as "secondary assessment", by which the resources available to deal with a situation are analysed. With regard to the assessment of own abilities, novice drivers display a tendency to overestimate their own driving skills.

*Demand component "Weighing up of risk"*
According to Deery (1999), the importance of the weighing up of risk derives from the fact that participation in motorised road traffic involves also self-imposed demands (e.g. own choice of speed and safe distance). The driver defines these demands by placing an assessment of the risk of damage or injury occurring in a given situation in relation to a self-assessment of his ability to act appropriately. In this context, Deery (1999) assigns a certain role to risk acceptance, which must be viewed in conjunction with motivational and personality-related aspects and represents an individual risk threshold which the driver is willing to accept before seeing a need to react.

*Demand component "Decision"*
The decision as to which behaviour is best able to contribute to hazard avoidance in a certain situation is a further important demand relating to traffic perception and hazard avoidance. In some situations, for example, emergency braking may be required to avert a hazard. Conversely, there are also conceivable situations in which sudden braking may

be inappropriate (e.g. the following vehicle is already very close behind). The decision taken by the driver is dependent on his subjective assessment of the risk entailed by the traffic situation. The decision is at the same time influenced by the repertoire of behaviour options available to the driver and the anticipated consequences of a particular choice of behaviour. In the context of the demand component “decision”, novice drivers tend, for example, to choose speeds which are too high and thus inappropriate in the current situation (Grattenthaler et al., 2009).

*Demand component “Action”*
This demand component covers actual execution of the driving behaviour selected from the available options. Sturzbecher et al. (2005) here speak of proper realisation. Actions thus refer to the level of vehicle operation and include psychomotor demands. Typical deficits displayed by novice drivers include, for example, the poor execution of appropriate steering movements (Grattenthaler et al., 2009).

*Summary*
The above presentation of a linear sequence of demands or structural components of traffic perception and hazard avoidance is the result of a content analysis approach and serves to simplify description. When it comes to the mastering of demands in real traffic, however, it can be assumed that certain processes will run concurrently: The acquisition and evaluation of information as a basis from which to derive action decisions, for example, is a continuously ongoing process; within this process, the individual structural components are closely interrelated. This conception is also consistent with the models described by Schlag (2009) and Deery (1999). Schlag (2009), for example, points to the immediate connections between the stages of detection, localisation and identification. For Deery (1999), on the other hand, the assessment aspects of risk perception, self-assessment of skills and risk acceptance stand in close relationship to each other.

To conclude, the choice of the term “traffic perception and hazard avoidance” to designate the described construct is to be founded. In English-language literature on the subject, the term “hazard perception” is often used. The various definitions found for this term, however, indicate that no common understanding of the associated content exists among scholars in the field: McKenna and Crick (1994b), for example, describe hazard perception as “the ability to respond to potentially dangerous traffic situations”, while Horswill and McKenna (2004) speak of “the ability to anticipate dangerous situations on the road ahead”. Both definitions contain an explicit reference to hazard situations. Other definitions of hazard perception, such as those of Mills et al. (1998; “the ability to ‘read the road’”) or Catchpole and Leadbeatter (2000; “the ability to observe the whole road scene, to identify and assess possible hazards in the traffic environment and to make timely responses to avoid or deal with the possible hazard”), by contrast, refer not only to the hazards, but also to road traffic in general. The latter definition, furthermore, goes beyond mere perception of the traffic environment to expressly include hazard avoidance.

It appears evident that the eight construct components elaborated above are not adequately represented with the term “hazard perception”, not least because this fails to cover the necessary action processes of hazard avoidance. Use of the term “hazard perception” also seems to be problematic from the pedagogical perspective, as the fundamental learning objective for novice drivers is only to a certain extent the recognition of existing hazards. According to Munsch (1973), the most significant learning objective is instead the ability to become active before a hazard actually develops. The terms “traffic sense” or “premonition” which are used by Munsch in this context emphasise the importance of timely perception of the situational cues which permit action to be taken “prior to danger”. The term “hazard perception”, on the other hand, focuses on a point in the development of a situation which has already progressed beyond the stage of “prior to danger”.

Finally, the term “hazard perception” suggests that hazards, from the driver's perspective, are manifested as concrete objects in the traffic environment (e.g. a child or a cyclist) to which the (additional) attribute “hazardous” can be assigned. This seems short-sighted, however, because hazards in road traffic generally result from the interactions of different participants (e.g. the dynamics of their behaviour, the direction of their motion or their adherence to rules) – and the driver's own person is here inextricable as a co-participant. For all these reasons, it was decided to use not the term “hazard perception”, but rather the construct designation “traffic perception and hazard avoidance” in the present report. This designation is intended to convey the actual demand placed on the driver: Constant observation of the whole traffic environment, in order to discover cues to potentially arising hazards and, after assessment of those cues, to avert hazardous situations at an early stage in their development by way of appropriate actions.

## 1.7 Approaches to the measurement of traffic perception and hazard avoidance

### 1.7.1 Methodical foundations for test development

The elaboration of a traffic perception test must observe the general methodical conventions applicable to the development of achievement tests. This chapter will thus first sketch the essential steps in test development (e.g. Moosbrugger & Kelava, 2008; Lienert & Raatz, 1998) for the context of a traffic perception test. These steps include not least proof that the test satisfies quality criteria of psychological testing[19]. Subsequently, the existing approaches for the measurement of traffic perception, criteria for format selection and promising operationalisation approaches are to be described.

*Classification of traffic perception tests*

A future traffic perception test – like the TDT (a highly standardised knowledge test) and PDT (a “work sample” in real traffic with systematic behaviour observation) – is a form of achievement test. Achievement tests are characterised in that they comprise tasks and problems which require the reproduction of knowledge and the demonstration of skills and concentration (Rost, 2004). In the case of a traffic perception test, the assessment of achievements relating to sensory, cognitive and possibly also motor abilities is placed in the foreground. Irrespective of the specific content-related demands, achievement tests demand “maximum behaviour” (Moosbrugger & Kelava, 2008).

It is usual to distinguish two types of achievement test, namely “speed tests” and “power tests”. Speed tests are characterised by the limited period of time which is allowed to complete the set tasks. The different capabilities of the individual candidates are expressed in the number of tasks which are solved correctly within the available time. In a power test, by contrast, the level of difficulty is varied such that, even with unlimited time available, not all candidates will manage to solve all the tasks correctly. Moosbrug-

[19] The quality criteria defined in classic test theory enable a test to be judged and founded according to its contextual and methodical quality. The principal quality criteria are considered to be “objectivity”, “reliability” and “validity”; secondary quality criteria are “standardisation”, “comparability”, “usefulness” and “economy” (Lienert & Raatz, 1998). The significance of these criteria in the context of driving licence testing has already been discussed in detail on numerous occasions (on the quality of the TDT: Bönninger & Sturzbecher, 2005; on the quality of the PDT: Sturzbecher, Mörl & Kaltenbaek, 2014) and is thus not to be expounded further at this point.

ger and Kelava (2008) point out that many achievement tests combine elements of both basic types. This is probably also the case for a future traffic perception test: On the one hand, the candidate's reaction time (e.g. the time before a hazard cue is recognised) is to be taken into account in the performance assessment. On the other hand, the demands to be met with regard to the identification of relevant situation attributes also permit differentiation according to a level of achievement (e.g. the number of hazard cues recognised). At the same time, a traffic perception test with relevance in the context of driver licensing must be classified as a criterion-referenced test (Amelang & Zielinski, 2002) which aims to assess the attainment of learning objectives, i.e. adequate competence in respect of traffic perception and hazard avoidance.

*Scope of validity and target group*

Before commencing test development, it is necessary to clarify not only the test purpose, but also the target group. In the case of a traffic perception test for use in the context of driver licensing, a very broad target group must be defined, e.g. it must include different age groups. The scope of content validity should permit determination of any safety-relevant novice-specific deficits, or else confirmation of adequate competence relating to traffic perception and hazard avoidance, wherein it can be assumed necessary to integrate different demand components (e.g. “observation”, “assessment”), as was already revealed clearly by the discussions in Chapter 2.2.3 above. The extent to which a claimed scope of validity is actually covered by the test must then be verified in corresponding validity studies.

*Attribute analysis and structural test design*

Content-oriented and psychological analysis of the attributes of interest is a prerequisite for the construction of test items (Lienert & Raatz, 1998). With regard to the development of a traffic perception test, both the already realised analysis of demand components (see Chapter 2.2.3) and existing empirical findings indicate that the attribute to be measured is a heterogeneous construct. A factor analysis study conducted by McKenna and Horswill (1997; cited after Wetton et al., 2011) to investigate candidate performances in four different hazard perception tests, each with a different content focus (“safe distance to the preceding vehicle”, “overtaking”, “safe distances to crossing and oncoming traffic” and “anticipation of hazards”), revealed factors which were independent of each other: Good performance in one area of the demands does not necessarily correlate with good performance in the other areas. It is not possible to measure such heterogeneous constructs by way of an homogeneous, unidimensional test. It is rather the case that the individual aspects must be assessed by way of suitable sub-tests or series of different test methods which, as an overall test battery, supply a differentiated performance profile to permit determination of an individual's performance. The length of the test, i.e. the number of tasks to be completed by the candidate, is greater for a heterogeneous construct, as a certain number of items is required for each sub-test. The test duration is extended accordingly.

*Test item construction: Test item types, instruction and answer formats*

The test items used in the current TDT are exclusively multiple-choice and gap-fill questions. For the area of traffic perception and hazard avoidance, however, alternative formats which require the localisation of hazardous objects or else behaviour decisions under time pressure, for example, appear promising. Each test item comprises an element serving to present the question (“instructions”) and an element to acquire the candidate's response. The various design possibilities for the instructions (including illustration of the underlying traffic situations) are referred to in the following as “instruction formats”; the different designs for answer templates are referred to as “answer formats” (Friedel, Weiße, Genschow, & Schmidt, 2011).

For the instruction formats of a traffic perception test, the stimulus material used to visualise the information to be detected by the candidate plays a particular role. It is possible to use either static images or dynamic situation presentations, and this stimulus material can take the form of either computer-generated visualisations or – as a costly alternative from the perspective of test economy – photographs of real traffic situations. Adequate empirical findings from which clear recommendations can be derived for a focus on a particular form of presentation are currently unknown. Such conclusions could only be drawn from a comparative analysis of the available design options. It would nevertheless remain uncertain whether these findings can be generalised or whether they are only valid for the particular contents of the stimulus material concerned.

Empirical studies conducted to date on the validity of test methods indicate that all the aforementioned design options yield valid stimulus material. Static photographs of real traffic situations, for example, were used as stimulus material by Vlakfeld (2011) and by Lyon, Borkenhagen, Scialfa, Deschênes and Horswill (2011); real video sequences have also been a frequent choice (e.g. McKenna & Crick, 1994b; Wetton, Hill & Horswill, 2011; Grayson & Sexton, 2002). A comparative study on differences between computer-generated scenarios and real-traffic video clips by Sümer, Ünal and Birdal (2007) suggests that the use of real video more frequently enables discrimination between experts and novices than is the case with computer-animated scenes. From the methodical perspective, however, this study fails to supply adequately sound evidence for the general superiority of real videos.[20] In the end, possible differences between video sequences and computer-animated scenes are to a large degree dependent on the perceived quality of the visualisation, although it must be said that technical advances in the field of computer animation are constantly narrowing the gap to real video material. This is corroborated by the study results of Malone, Biermann, Brünken and Buch (2012), which show that a valid discrimination between experts and novices can be achieved with computer-generated stimulus material.

The question which arises in connection with video-based driving scenarios is whether relevant traffic situations should be staged, or whether actually encountered situations should be used. According to Wetton et al. (2011), staged traffic situations possess certain disadvantages compared to genuine, un-staged situations: Firstly, it is considered problematic that the traffic situations must be designed by experts with driving experience, and that precisely this driving experience influences the design process. As a result, it would not be surprising if experienced drivers were to recognise the incorporated hazards better than inexperienced drivers. Secondly, the authors are wary of a possible limitation of the content validity, as the staged driving scenarios may not actually occur in real traffic on the road; in the context of driving licence testing, this could lead to objections from test candidates. These reservations seems plausible, but even in the case of a real drive, it is left to experts to determine which hazard cues are relevant, and how or how quickly the driver is expected to react.

In Great Britain, actual traffic situations are used to prepare a traffic perception test. However, the original stimulus material in the form of real video is here reproduced –

[20] The determined differences are based on a comparison of absolute figures and are actually reversed when these figures are placed in relation to the total number of test items analysed in each case: Although 9 (out of a total of 27) real videos, but only 4 (out of a total of 10) computer-animated sequences produced significant differences between the results achieved by experts and novices, relativisation to the totals of each format reveals a higher frequency of discrimination for the computer-generated sequences (4 out of 10, i.e. 40%) than for the real videos (9 out of 27, i.e. approx. 33%).

maintaining all situation attributes of the scene – using computer-generated imagery (Wedell-Hall, 2013). Through this combination of different production techniques, it is possible to realise the test economy benefits of computer-generated stimulus material and at the same time to answer the aforementioned methodical reservations concerning the creation of such stimulus material.

*Trial studies and task analysis*

A test item can be deemed valid in terms of the attribute to be measured if those persons among whom the attribute concerned is most evident are found to display the expected performance more frequently than others among whom this attribute is less pronounced (Lienert & Raatz, 1998). One practicable approach by which to test this requirement is the so-called "known groups method", which evaluates the performances of persons known to display a certain attribute to a greater or lesser degree in order to determine whether these known differences are reflected in their handling of the test items concerned. Applied to the case of traffic perception and hazard avoidance in road traffic, the attribute chosen to define the two groups is usually the scope of driving experience (e.g. the reported amount of driving done in kilometres, the duration of driving licence possession in years). Persons with little driving experience are assigned to the group of "novice drivers", while those with extensive experience are classified as "experts", although further graduated differentiation may well be applied in both groups, as appropriate (e.g. professional drivers, driving test examiners, driving instructors or police officers as persons with especially well developed driving expertise; participants in a training measure as further qualified novices).

Where tests comprise video sequences in which hazard cues must be recognised and the further development of a given situation must be anticipated, expert-novice comparisons reveal differences in response latency, which usually lie within the range of several hundred milliseconds (e.g. McKenna & Crick, 1994b; Wetton, Hill & Horswill, 2011). Such performance differences provided a basis for the task analysis conducted by McKenna and Crick (1994b) within the framework of a study serving to develop a corresponding test. Taking an original pool of 35 videos in which a developing hazard was to be indicated by pressing a button, they used the analysis of (co-)variance[21] to select those videos where the expert and novice drivers displayed the most evident differences in response latency. To ensure that these differences between the expert and novice drivers were not merely an expression of differences in the time taken to actually press the button, the reaction times of the participants were also recorded so that their influence could be controlled statistically. Only those test items where the performance differences between expert and novice drivers remained after controlling for the reaction time were finally chosen. When applying this methodical approach to test development, it must be taken into account that extreme groups are being compared to identify differences in performance between experts and non-experts. In the case of a test method specific to novice drivers, the performance differences which can be determined will presumably be more subtle, because abilities relating to traffic perception are overall less pronounced in this target group.

[21] In analysis of variance, group mean values are compared (here the average response latency of expert and novice drivers with regard to the individual test items). To test the significance, the overall variance, the variance between the groups and the variance of the values within each group are considered. The analysis of covariance also permits the influence of a variable which is deemed irrelevant for the study in hand to be excluded (here the individual speed of reaction to a given stimulus).

*Test construction and reliability studies*

To construct a test form appropriate for use as a test method in novice driver preparation, it is necessary to select a certain number of suitable test items and to compose corresponding task sets, taking into account the results of the task analysis. The creation of parallel task sets is also meaningful from the methodical perspective, as it permits the reliability of the test to be verified using the parallel-forms method (see above). Parallel tests are at the same time useful from the point of view of test didactics, because a possible repeat test can be based on a new set of tasks, rather than those which are already known to the candidate. The methods used to determine the reliability, and similarly the demands set in respect of the reliability coefficient to be achieved, are dependent on the construct to be measured. In the case of abilities relating to traffic perception and hazard avoidance, it can be assumed that a learning effect follows from driving experience. Accordingly, a reduced correlation can be expected between the separate instances when measurements are repeated to determine the test-retest reliability. It thus seems more expedient to use the parallel-forms or split-half method and to determine the internal consistency (which should always be found to be high).

Generally speaking, reliability coefficients between .80 and .90 are viewed as moderate, while values over .90 are considered high (Weise, 1975, cited after Bortz & Döring, 2006; Fisseni, 1990). For a test with dynamic situation presentation, McKenna und Crick (1994b) report good consistency between two parallel forms (with a correlation of .85 for a test mean recorded as the response latency in milliseconds). They also calculated the inter-item correlation as a further measure of reliability, and here determined a medium degree of internal consistency for both test forms (Cronbach's alpha = .78 and .82). Wetton et al. (2011) report similar inter-item correlations ranging from .73 to .81 for the reliability of a total of four parallel test forms (with 15 items each). Higher values are presented by Grayson and Sexton (2002), who calculated reliability coefficients between .82 and .86 for four parallel test forms (each comprising 13 video sequences with 22 hazard cues to be recognised). Regarding the reliability coefficients of hazard perception tests, Horswill and McKenna (2004) establish that, overall, the values obtained by different methods vary considerably and are in some cases also very low (e.g. Cronbach's alpha = .48; Pelz & Krupat, 1974). A possible explanation for these differences is seen in the partly homogeneous and partly rather heterogeneous nature of the construct to be measured when it comes to the specific test demands and the design of the stimulus material.

*Determination of cut-off points*

For a test method which serves performance assessment and is intended to judge an attained level of learning or experience, the accuracy of the classification as either a “master” or “non-master” is a relevant criterion for the quality of the method. Alongside correct classifications which lead to masters passing the test (“true positive classification”) and non-masters failing (“true negative classification“), two forms of false classification are possible when the final dichotomous test decision is determined: Firstly, non-masters could nevertheless pass the test („false positive cases”), and secondly, candidates who are actually masters could fail the test (“false negative cases”). The terms “sensitivity” and “specificity” are used to express the precision of a test: A traffic perception test is deemed sensitive, if it detects all those persons who are not adequately prepared; it is deemed specific, if only those persons who are not adequately prepared are denied participation in motorised road traffic.

The items chosen to construct the test should be those which promise the highest possible classification accuracy; this can be tested empirically through recourse to the known-groups method (see above). Lyon et al. (2011), for example, report that, using a test method which they had developed with 21 image-based test items, correct assignment to the group of either experienced drivers or inexperienced novices was possible for 78 per cent of the study participants (N=51) on the basis of their indicating a hazard using a touchscreen and measurement of the individual response latency. The sensitivity was found to be 84 per cent, the specificity 73 per cent. For a different method with four test items based on dynamic situation presentations, McKenna and Crick (1994b) report that the differences in response latency of the study participants permit correct classification as an expert or novice in 100 per cent of the cases.

The sensitivity and specificity of an achievement test method are directly dependent on the cut-off point defined as a basis for the classification decision. If the chosen value is too low, unprepared candidates may also be able to pass the test, i.e. the sensitivity is reduced. A high cut-off point would increase the sensitivity, but that only at the cost of reduced specificity: More non-masters would be identified, but at the same time, an increasing number of masters would be falsely classified as non-masters. To determine an ideal cut-off point, Wetton et al. (2011) applied the methods of signal detection theory (Nett & Frings, 2014) and analysed the relationship between “true positive cases” and “false positive cases” for all possible cut-off points by plotting a receiver operating characteristic (ROC). They then chose the cut-off point which achieved the best possible discrimination between experienced and novice drivers. For the development of a traffic perception test to be used in driving licence testing, it must be remembered that the cut-off point is intended to distinguish between trained and untrained candidates; the actual level of performance which can be expected in the target group at the time of testing is to a large extent dependent on the effectiveness of the training possibilities offered.

*Test validation by way of external criteria*
The criterion-referenced validity is to be considered the most important aspect of test validation (Lienert & Raatz, 1998). This form of validity refers to the correlations between test performance and a relevant “external” criterion. In the case of a traffic perception test which is intended to judge the adequate acquisition of competence relating to traffic perception and hazard avoidance, one particularly relevant external criterion is accident involvement in real traffic: If poor test performances correlate with more frequent accident involvement, the specification of a minimum required level of achievement could (for the time being) systematically exclude inadequately prepared test candidates from participation in road traffic and thus prevent accidents. In this specific context, however, determination of the predictive validity is relatively complex, due to the time lag between the assessment of test performance and the further surveying of accident involvement or retrieval of corresponding information from accident databases. The necessity to ensure paired samples (i.e. the same persons must participate at each stage of the study over an extended period of time), in particular, places high burdens on organisation. A far less complex alternative would be to ask candidates to report on previous (near-)accidents at the time of test participation (“retrospective accident reporting”). If the persons displaying poor performance were also to report more frequent involvement in critical situations, this could be taken as a sign of validity.

Despite the fact that accidents are evidently of high practical relevance, the validity of results from (prospective and retrospective) accident analyses remains subject to certain reservations from the methodical perspective. Verification of the predictive validity by way of accident analyses, for example, appears very difficult in connection with driving licence testing, because the variance in the data is limited (Bönninger & Sturzbecher, 2005): Firstly, inadequately prepared test candidates and novice drivers are automatically excluded from later accident analyses because they are denied the entitlement to participate in road traffic. At the same time, successful test candidates are assumed to possess adequate competence and thus – ideally – will also not become involved in accidents.

Furthermore, fundamental questions arise with regard to suitable and feasible means to operationalise accident rates. Information on the duration of driving licence possession, for example, is generally more readily available than information on a driver's mileage during this period. The latter aspect, however, would actually produce sounder results, as accident involvement is dependent on the scope of exposure to the corresponding risk. In other words, the probability of accident involvement increases with the distance driven over a given time period. Further methodical problems are summarised by Horswill, Anstey, Hatherly and Wood (2010):

(1) Accidents are very rare occurrences and thus a very large sample must be recruited if it is to contain a sufficient number of persons with prior accident involvement.

(2) The accurate recording of accidents is very difficult. The problem which arises in connection with self-reported accidents is that drivers tend to forget up to 30 per cent of the relevant incidents which occur in any one year. Police statistics, on the other hand, usually cover only relatively serious accidents, and were furthermore found to correlate with self-reported accidents to only a minor extent.

(3) Accidents can be attributed to a multitude of factors and are thus not necessarily to be associated with poor driving ability.

This chapter considered the methodical steps to be taken in test development. Building upon this analysis and the description of the construct “traffic perception and hazard avoidance” developed in Chapter 2.2.3, approaches to construct measurement are now to be presented and discussed with regard to their suitability in the context of driver training and driving licence testing. First of all, forms of competence measurement which are already in use within the framework of regular driving licence testing are to be described. Subsequently, a number of operationalisation approaches derived from experimental studies are to be investigated. It must here be taken into account, however, that experimental research – although by all means a source of rateable empirical findings – is frequently conducted under laboratory conditions which cannot be reproduced without considerable effort in a mass test procedure like a driving licence test (e.g. use of eye-movement trackers or personality questionnaires). Even so, systematic combination of the various operationalisation approaches and incorporation of the empirical findings promise a fertile starting point for the development of innovative instruction formats to assess traffic perception and hazard avoidance.

## 1.7.2 Traffic perception tests in international driving licence testing

Traffic perception tests are currently implemented in two European countries – Great Britain and the Netherlands – and in five Australian states – New South Wales, Queensland, Victoria, Western Australia and South Australia. In the following, country-specific design variants of the traffic perception test are described together with evaluation results pertaining to the method quality, where available.

*Great Britain*

Since 14th November 2002, driving licence applicants in Great Britain must complete a “Hazard Perception Test”, which comprises 14 one-minute dynamic driving scenarios viewed from the perspective of the driver. The scenarios presented to the candidate are recordings from real drives; in future, they are to be replaced with computer-generated sequences, although all the original content is to be maintained (see above). The task in each case is to identify potential hazards which demand an appropriate reaction and to indicate the hazard as quickly as possible by clicking with the mouse. The test performance of the candidate is assessed on the basis of the response latency. Depending on how quickly the candidate responds to the emerging hazard (e.g. a pedestrian on the road), between 0 and 5 points are awarded: If the hazard is not recognised within the expected time, the candidate receives 0 points; if the hazard is recognised immediately, 5 points are awarded. There are 15 hazards to be identified in total (one of the videos contains two hazards) and the maximum attainable score is thus 75 points; a score of 44 points is necessary to pass the test.

The validity of the hazard perception test was investigated within the framework of a comprehensive long-term study (Wells, Tong, Sexton, Grayson & Jones, 2008). Over a four-year period starting in November 2001, i.e. before introduction of the hazard perception test, learner drivers who were about to take their driving test were asked to report about their driver training by way of a questionnaire (e.g. duration of training, number of driving lessons). After passing the test and commencing their solo driving careers, the study participants were contacted once more for further surveys at specified intervals (6, 12, 24 and 36 months after receipt of the first questionnaire). One central objective of the study was to determine the extent to which implementation of the hazard perception test, i.e. the introduction of an additional, innovative test form in driving licence testing, influences accident involvement during later solo participation in road traffic. Evaluation of the self-reported accidents in which survey respondents were involved showed that the probability of an accident on a "non-low-speed public road"[22] was 3 per cent lower where the driver had passed a hazard perception test.

Wells et al. (2008) also analysed whether performance differences in the hazard perception test displayed any correlations to later accident involvement in real traffic ("predictive validity"). To this end, the test performance (total score) was taken to represent the hazard perception ability of the candidate. The analysis revealed a statistically significant inverse correlation between the multiply graduated variable "hazard perception score" and the probability of accident involvement during the first year of solo driving: Novice drivers who had passed the hazard perception test with a high score displayed a lesser accident risk compared to those who had passed with a lower score. The difference in accident involvement for accidents which occurred on non-low-speed public roads and in which the driver considered himself to be (partly) responsible (i.e. partial blame was acknowledged in the questionnaire) amounted to 4.5 per cent. The authors attribute this result to the greater effect which is naturally to be expected in respect of this type of accident where a driver displays better hazard perception.

Mason (2003) sought to verify the validity of the hazard perception test by considering performance in the driving test as an external criterion. The expected correlations between hazard perception performance and passing or failing of the driving test could not be proven. In certain demand situations during the driving test (e.g. "pedestrians crossing the road", "use of mirrors", "signalling"), however, candidates who achieved high scores in the hazard perception test did indeed display better performance. The author interpreted the determined correlations as an indication of the criterion validity of the hazard perception test, even though the findings of the study seem not to be wholly conclusive.

[22] The authors fail to provide a more precise definition of the term "non-low-speed public road", e.g. by specifying a speed or speed limit. In Great Britain, the maximum permissible speed within built-up areas is 30 mph (48 km/h). Outside built-up areas, the speed limit for cars is 60 mph (97 km/h) on roads without physical separation between the traffic moving in opposing directions or 70 mph (113 km/h) on dual carriageways and motorways.

*Netherlands*

In the Netherlands, driving licence applicants have been required to take a traffic perception test since 2009. The candidate is shown 25 photographs depicting traffic situations from the driver's perspective (with the vehicle mirrors, turn indicators and speedometer visible). Within a given response time of eight seconds, the candidate must decide which of three possible behaviour options “Apply the brakes”, “Take foot off the accelerator” or “Do nothing” is most appropriate in the depicted situation; these same three options are offered for every test item. The action alternatives correspond to two different types of hazard which may be depicted. So-called “acute hazards” (e.g. a pedestrian who wants to catch a bus waiting on the other side of the road and crosses without paying attention to the traffic) require the reaction “Apply the brakes”. The reaction “Take foot off the accelerator”, on the other hand, would be appropriate in the case of a “latent hazard” (e.g. a bus standing at a bus stop, and thus the possibility that passengers could alight and wish cross the road, even though no persons are actually visible in the depicted scene). To pass the test, at least 12 of the 25 test items must be answered correctly.

To test the validity of the Dutch traffic perception test, Vlakveld (2011) compared the test performances of novices (learner drivers shortly before they passed the driving test) with those of experts (drivers with 18 months of driving experience and drivers with at least 10 years of driving experience). It was shown that the experts achieved significantly better test results than the novice drivers. There were no significant differences, however, between the groups of experts with 18 months and 10 years of driving experience. In the expert group with 18 months of driving experience, an additional analysis was performed to consider a possible correlation between performance in the traffic perception test and the number of self-reported accidents, for which purpose differences in the amount of driving done were similarly acquired and taken into account in the evaluation as a factor which could influence accident frequency. The result of this analysis was that, on average, drivers with an accident-free driving record achieved significantly better test results than drivers who reported accident involvement. These findings suggest that safety-relevant performance differences can be measured with the traffic perception test used. As the driver groups were compared solely on the basis of their overall scores, however, it is not evident whether (and, if so, what) differences exist with regard to the individual demand aspects (differentiation of latent hazards, acute hazards and non-hazardous situations).

*New South Wales (Australia)*

In the Australian state of New South Wales, two touchscreen-based traffic perception tests are taken at an interval of 24 months (see Chapter 2.4). In both cases, visual and acoustic signals are given to the candidate to confirm that his reaction has been recorded. Each test begins with two practice questions, which serve to ensure that the candidate has properly understood the test requirements.

The first test comprises 15 items, for which traffic situations are presented from the driver's perspective in the form of 30-second real-life videos. The candidate is also able to see the speed of his own vehicle and the activation of the vehicle's turn indicators during each video clip. Three content-related demands are addressed: Observance of the necessary safe distance to other vehicles, selection of an appropriate moment to perform a specified driving manoeuvre and the identification of hazards. The candidate is asked to touch the screen as soon as he deems it safe to perform a certain action. Before each video sequence, a text instruction appears on the screen to briefly describe the subsequent traffic situation (e.g. “You are travelling along a two-way road in a 60 km/h speed zone and wish to keep driving straight ahead”), together with specification of the subsequently required behaviour (e.g. “Touch the screen when you would slow down”). The requirement is thus not to select the correct behaviour response from a multitude of possible options, but rather to decide when the execution of a given action is reasonable. Each question involves one single action. Possible actions are “Slow down”, “Overtake” or “Turning/crossing at a junction”. The test demands are based on the five most common types of accident involving novice drivers in New South Wales.

The second hazard perception test is similarly a touchscreen-based test in which the candidate must indicate when it is safe to perform a specified action (e.g. turning at a junction). The demands addressed are the same as for the first test. Unlike the first test, however, the video clips do not necessarily contain only one instance of the required action; there may be several occasions on which the given action can be performed. Furthermore, only ten video clips are presented, though they are at the same time longer than those shown in the first traffic perception test. After completing all the required test items, the novice driver receives feedback on whether or not the test was passed; a concrete score is not communicated. If the test is failed, the candidate also receives pointers to the areas in which the demands were not met. Where appropriate, feedback on remaining competence deficits is similarly provided to successful candidates. There have apparently been no empirical studies to date on the methodical quality of the test procedure.

*Queensland (Australia)*

In Queensland, driving licence applicants must complete an online test (“Queensland Transport Hazard Perception Test” – “QT-HPT”), for which they are presented video clips of real traffic scenarios in which so-called “traffic conflicts” could arise. A traffic conflict is described as a situation in which the candidate's own vehicle would collide with another road user (e.g. another motor vehicle, a pedestrian or a cyclist) if certain actions are not taken (e.g. slow down or change direction). The task for the novice driver is to identify conflict situations as quickly as possible and then to mark the position of the hazardous object concerned by clicking with the mouse. Each driving scenario contains a potential traffic conflict; the candidate's reaction is marked by way of a yellow circle, which at the same time serves to confirm that the click has been recorded. The test result is determined on the basis of reaction times and the positions of the mouse clicks. At the end of the test, the candidate is informed as to whether the test was passed. In addition, feedback is given on the individual test performance. Especially if the test is failed, this includes detailed advice on the areas in which improvement is necessary.

Wetton, Hill and Horswill (2011) compared the results of experienced drivers (n=56; driving licence held for at least 15 years and an average of 19,671 km driving per year) and novice drivers (n=94; driving licence held for 4.2 months on average and an average of 3,439 km driving per year) in a preliminary version of the QT-HPT. To this end, each of the study participants completed 91 test items. It was shown that the mean response latency of the experienced drivers (6.321 seconds) was lower than that of their novice counterparts (7.069 seconds). The differences in response latency over all items were statistically significant. For the final test form, the original scope was reduced to those 60 items for which the differences in response latency between the two groups were especially obvious. After this selection process, the mean difference in response latency for the remaining items was 1.022 seconds. The authors illustrated the practical significance of this apparently small difference by drawing a reference to the circumstances of actual traffic participation: Already at the relatively low driving speed of 60 km/h, this delay equates to an additional distance of approx. 17 metres travelled before response to a hazard cue. The validated pool of test items served as the basis for construction of a total of four parallel tests.

*Victoria (Australia)*
In Victoria, driving licence applicants are required to take a hazard perception test before being allowed to attempt the practical driving test. The hazard perception test comprises 28 items, each of which is illustrated by a 30-second video of a traffic scenario viewed from the driver's perspective. Before each video, the required driving action ("slow down", "overtake", "start to turn") is specified in corresponding text instructions. The task for the candidate is to decide when – during the course of the depicted scenario – the previously specified action can be performed safely. The test result is determined according to whether the candidate reacts appropriately within a defined time window. There are certain test items, however, where the specified action should never be taken ("no response items").

In 1999, Congdon and Cavallo validated an earlier form of the hazard perception test; their study results were primarily statements on the predictive test validity. To this end, the accident data of almost 100,000 novice drivers were collected within the framework of a census and evaluated for a period of around 18 months from the commencement of solo driving. According to police records, persons from the study population were involved in approximately 2,300 accidents which resulted in injury to persons during this period. The authors distinguished three categories of accident severity, which were then taken into account as dependent or criterion variables: (1) "Fatal accidents", (2) "Serious casualty accidents" and (3) "Casualty accidents". The factors chosen as independent variables or predictors included the score achieved in the hazard perception test and the period of driving licence possession. By way of regression analyses, it was then determined whether and, if so, to what extent possible accident involvement can be predicted from these variables ("predictive validity"). The results showed that the probability of involvement in a fatal accident or an accident with serious casualties is one to three times higher for persons with a relatively poor result in the HPT compared to persons with relatively good results. The HPT score was found to be a meaningful predictor for fatal accidents for a period of up to 12 months from the granting of a driving licence, but was then no longer significant during the remainder of the 18-month period.

The findings of Congdon and Cavallo (1999) were taken into account in the development and implementation of the presently used version of the hazard perception test (see above) in Victoria. A renewed evaluation of the further developed test method was performed in 2001 by Catchpole, Congdon and Leadbeatter (2001). To this end, a total of 363 drivers were divided into three groups with different scopes of driving experience: (1) Driving licence applicants, (2) Drivers with 2 years of driving experience, and (3) Driving experts with at least 15 years of driving experience, at least 300,000 miles of driving done and only few accidents. Comparison of the test results revealed the expected differences in the sense that better test performances correlated with greater driving experience. These differences, however, were not statistically significant.

*South Australia (Australia)*
In South Australia, a hazard perception test in which the driving licence applicant must demonstrate that he is able to identify and react appropriately to potentially hazardous situations was introduced on 31st October 2005. The test demands refer, for example, to keeping a safe distance from other vehicles, selecting safe gaps and scanning the surroundings for hazards in front of, behind or to the side of one's own vehicle. Each test begins with two practice questions. Subsequently, the candidate must complete 15 test items chosen at random from a larger pool of questions. Each test item consists of a video clip of real traffic situations viewed from the driver's perspective. A dashboard shows the speed at which the vehicle is travelling and the state of the turn indicators. Before the video clip starts, the coming traffic situation is explained more closely and a still picture

of the starting point is presented for a few seconds. At the same time, instructions are given as to the action (“slow down”, “overtake” or “cross the intersection”) which is to be taken during the course of the depicted scenario. The candidate must then touch the screen when he thinks the specified action can be taken safely. If the specified action is not considered to be safely possible at any time during the video clip, then the candidate can also choose not to touch the screen. Only one response is required in each case; the screen flashes briefly to indicate that a response was recorded. In addition, the candidate is asked to confirm whether or not a response was given (the two buttons “You touched the screen” and “You did not touch the screen” are displayed at the end of each video clip).

Immediate feedback is provided only in conjunction with the initial practice questions. The candidate is told whether the displayed response was “good”, “in need of improvement” or “unsafe”, and is also offered the opportunity to try the same test item again. Once the actual test begins, however, no immediate feedback is given and it is not possible to repeat individual test items. At the end of the test, the examiner informs the candidate as to whether he has passed or failed. Feedback is also given on aspects of the test which could be improved. In the case of a negative result, the candidate may resit the test immediately, provided the test centre is able to offer a corresponding vacancy. The transport authority website[23] presents a number of exercises to help driving licence applicants to prepare for the test. There have been no empirical studies relating to the methodical quality of the test procedure implemented in South Australia.

*Western Australia (Australia)*
In Western Australia, the hazard perception test comprises 28 test items with video clips viewed from the driver's perspective; the vehicle speed and the state of the turn indicators are also displayed. Before each video clip, the candidate is given text-based instructions to describe the initial situation (e.g. “You are driving straight ahead”, “You are waiting for an opportunity to turn right”). One of four driving tasks is then specified for the coming scenario: “Slow down”, “Overtake”, “Make the turn” or “Move off”. The task for the candidate is to determine a suitable moment to perform the required action and to indicate this by clicking with the mouse, though there may also be video clips during which the specified action is never appropriate. Four practice items can be answered as familiarisation before the actual test begins. No studies exist on the methodical quality of this traffic perception test.

Viewed overall, it can be seen that the various methods used for the testing of traffic perception and hazard avoidance have a great deal in common, but nevertheless display conspicuous differences. This refers not least also to the questions of empirical verification and the availability of data on the quality of the individual methods from the perspective of test psychology. The empirical studies conducted to date (e.g. Wetton et al., 2011; Catchpole et al., 2001; Vlakveld, 2011) seem to confirm the corresponding quality of the method concerned. Table 2 below compares the different test methods according to the information available on instruction and answer formats, the parameters used to measure performance and the criteria applied to determine the test result.

[23] http://www.mylicence.sa.gov.au/hazard-perception-test (retrieved on 01.04.2014)

**Tab. 2: Operationalisation of traffic perception and hazard avoidance abilities within the framework of international driving licence testing**

| Country | Instruction format | Answer format | Performance parameter | Test assessment |
|---|---|---|---|---|
| GB | Dynamic situation presentation from the driver's perspective (computer-generated, presentation time 60 seconds) | Random mouse-click as indicator of a non-specific need for action | Response latency between first hazard cue and action execution | 14 test items requiring 15 computer responses; 0 to 5 points are awarded for each response; at least 44 of the total of 75 points are required to pass the test. |
| NL | Static situation presentation (real photos, presentation time 8 seconds) | Selection of the correct action from three options | Appropriateness of the selected action in the situation | 25 test items requiring 25 computer responses; at least 12 of the 25 items must be answered correctly to pass the test |
| NSW | Dynamic situation presentation (real videos, presentation time 30 seconds) viewed from the driver's perspective | Random touching of the screen as indication for execution of a specific action | Appropriateness of the moment of action execution in the situation | 1st test: 15 test items; maximum one response required per test item.<br>2nd test: 10 test items; a test item may require more than one response in some cases. Pass criterion not known. |
| QLD | Dynamic situation presentation (real videos) | Targeted mouse-click on the hazard-relevant aspect of the situation | Response latency between first hazard cue and action execution | 15 test items;<br>pass criterion not known. |
| VIC | Dynamic situation presentation (real videos, presentation time 30 seconds) viewed from the driver's perspective | Random mouse-click as indication for the moment of execution of a specific action | Appropriateness of the moment of action execution in the situation | 28 test items; a test item may require no response in some cases; at least 54 per cent of the best possible score must be reached to pass the test. |
| SA | Dynamic situation presentation (real videos) viewed from the driver's perspective | Random touching of the screen as indication for the moment of execution of a specific action | Appropriateness of the moment of action execution in the situation | 15 test items; a test item may require no response in some cases; pass criterion not known. |
| WA | Dynamic situation presentation (real videos) viewed from the driver's perspective | Random mouse-click as indication for the moment of execution of a specific action | Appropriateness of the moment of action execution in the situation | 28 test items; a test item may require no response in some cases; pass criterion not known. |

The tabular comparison shows that dynamic situation presentations are the most commonly chosen means to visualise traffic situations. This seems only natural, as traffic-related demands are especially dependent on dynamic developments in the traffic environment. Nevertheless, empirical studies on the importance of dynamic or static situation presentations indicate that expert and novice drivers also achieve different results when the test items are based on static images (e.g. Vlakfeld, 2011). It is furthermore shown that all visualisations presented to the test candidates assume the driver's perspective. Compared to the alternatives (e.g. bird's-eye view, perspective of other road users), this perspective possesses high face validity. For the development of learning and test media, however, the inclusion of other perspectives could well be meaningful (see Chapter 3), not least because perspective-taking is an important prerequisite for anticipation of the behaviour of other road users.

Regarding the answer formats and parameters used to measure performance, the (cognitive) test demands of the compared methods can essentially be divided into the following three types:

- Identification of situation attributes which require an otherwise unspecified action ("anticipation latency")
- Identification of situation attributes which necessitate the selection of a specific action ("action selection")
- Identification of situation attributes which indicate that it is safe to perform a specific action ("action timing").

The methods used in Queensland and Great Britain are based on the same performance parameter, namely the response latency between the first hazard cue and execution of the action. The test procedures in New South Wales, Western Australia and South Australia are similarly comparable, but instead require specification of when it is safe to perform a defined action. Wetton et al. (2011) justified the preference for response latency as a performance parameter in Queensland with their belief that test items which target appropriate action timing would reflect the individual readiness to take risks rather than the ability to recognise hazardous situations. The recording of risk acceptance by way of an achievement test is furthermore susceptible to manipulation, as the candidate could consciously modify his behaviour in the test situation, i.e. he could display a more safety-oriented behaviour than he would normally in a corresponding driving situation. In Queensland, the candidate is required to localise the hazard cues, which ensures that random clicking is less likely to be recorded as a correct response.

### 1.7.3 Approaches for the measurement of traffic perception and hazard avoidance in experimental studies

Beyond the analysis of those test methods which are already implemented in international practice, a number of experimental studies relating to the measurement of traffic perception and hazard avoidance were identified and evaluated within the framework of the present innovation report, paying particular attention to the chosen operationalisation approaches. Recourse to these studies serves to establish a broader, empirically founded basis for the development of suitable question formats for a traffic perception test. The operationalisation approaches found are outlined in the following sections, together with indications of the extent to which they were able to distinguish the performances of experienced and novice drivers.

*Change detection*

The research paradigm of "change detection" has a long tradition in basic perceptual psychology research (Rensink, 2002). In a study on the measurement of hazard detection competence, Wetton et al. (2011) chose this paradigm to reveal differences in the abilities of experts and novices to detect hazard cues. The test participants were presented a series of static images, wherein the task was to compare two ostensibly identical images, which differed only in the presence or absence of a potential hazard (e.g. a pedestrian standing at the kerbside), and to touch the screen at the place where a difference was seen (see example in Fig. 11). This form of operationalisation – at an elementary level of perception – was intended to identify the competence deficits of novice drivers.

**Fig. 11: Example for a test item based on change detection**

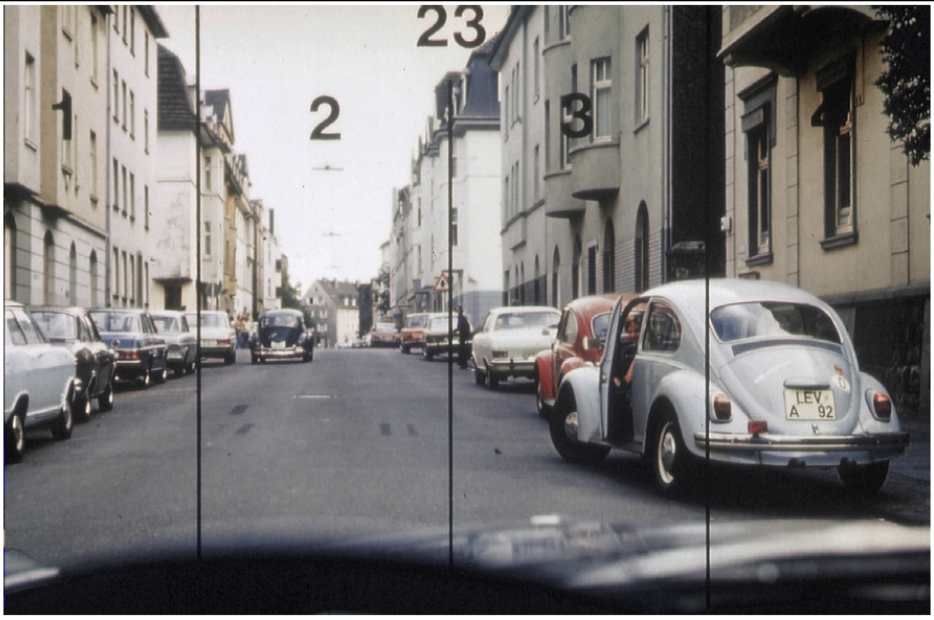
The parameters used to assess performance were the reaction time and the number of correctly identified differences. It was found that the results achieved by novice drivers were not poorer than those of experienced drivers. In fact, the reaction times recorded for the novice drivers were actually faster. On the basis of these results, the authors concluded that the change detection demand is not suitable to distinguish between experienced and novice drivers (Wetton et al., 2011). This finding is corroborated by accident analyses, which indicate that detection ability is above all age-dependent and decreases with increasing age (White, 2006). At the same time, accident analyses illustrate the high accident relevance of detection: In the case of accidents between cars and cyclists or motorcyclists, in particular, the cyclist or motorcyclist is sometimes overlooked as a road user, despite being in the car driver's field of vision. One of the factors put forward in the discussion of possible explanations for this “overlooking” of other road users is “change blindness”, i.e. the failure to recognise a relevant change in the visible environment (e.g. because this change is not sufficiently salient). A second possibility which is mentioned is “inattentional blindness”, i.e. the failure to recognise a relevant situation attribute (e.g. because attention is concentrated on other attributes).

Both the study results of Wetton et al. (2011) and the findings from accident analyses thus suggest that weaknesses in terms of change detection do not represent a novice-specific deficit which can be rectified by expanding driving experience. Accordingly, operationalisation variants which focus on the recording of change detection cannot be expected to contribute to a valid determination of performance differences between novice drivers. At first sight, this conclusion seems to stand in contradiction to the model concept of Grayson et al. (2003; see Chapter 2.2.2), which describes “hazard detection” as an important factor influencing novice accident risk. It must be noted in this context, however, that the method used by Grayson et al. to measure “hazard detection” did not comprise elementary change detection tasks; the chosen procedure (hazard perception test, appraisal of the traffic situation, spatial deductions) also covered demands which extend beyond mere hazard detection. The extent to which a component “hazard detection” can be deemed important in test development is thus also dependent on how narrowly or broadly this term is defined.

The assumption that systematic performance differences can only be revealed where “hazard detection” is understood to mean more than mere change detection can also be derived from the findings of Huestegge, Skottke, Anders, Müsseler and Debus (2010): Their eye-tracking measurements show that novices do not actually discover immediate hazard cues less quickly than experts, though they do take longer to indicate this detection. When developing new test formats, therefore, it is advisable to record not only whether a certain situation attribute is detected, but also whether it is classified as hazardous by the candidate.

*Hazard localisation*

An “audiovisual test system” was developed at TÜV Rheinland in the 1970s (Hampel, 1977a). This innovative test approach incorporated items for which the test demands were presented in the form of slide projections and voice recordings (“instruction format”). The candidate was required to identify and specify certain sections of the static images in which hazard-relevant information cues were to be seen (“answer format”; see Fig. 12).




**“Where do you see a hazard in this scene?”**

(Source: Cologne AV series, TÜV Rheinland; in Hampel, Sturzbecher et al. 2009)

**Fig. 12: Item from the audiovisual test system of TÜV Rheinland**

Empirical findings relating to the described test format can be found in the study of a convenience sample within the framework of the International Transport Exhibition in Hamburg in 1979 (Hampel, 1979). The results showed that participants with driving experience tended to solve the test items better than those without driving experience.

*Visual search strategies*

To investigate visual search strategies, Velichkovsky, Rothert, Kopf, Dornhöfer and Joos (2002) recorded the eye movements of study participants as they completed driving tasks on a simulator. The results indicate that experienced drivers employ specific gaze patterns to search for potential and direct hazards.

Chapman and Underwood (1998) used eye-tracking methods to analyse the gaze patterns of experts and novices by way of their eye movements during 20 to 70-second video clips. Each of these clips contained at least one hazardous situation (e.g. a braking vehicle); a variety of rural, urban and suburban environments was included. The response latency was measured by requiring the test person to press a button when a hazard appeared. The results showed that the periods of fixation were shorter for experts than for novices. One possible explanation, from the point of view of the authors, is that experts have already developed cognitive schemata for certain traffic situations and are thus able to interpret the fixated situation attributes more quickly. Furthermore, the fixation periods of the experts were dependent on the environment: They were longest in rural environments and shortest in urban situations. This supports the assumption that eye movement frequency increases with increasing situation complexity, and the fixation on individual objects is correspondingly shorter. Contrary to earlier findings (cf. Deery, 1999, in Chapter 2.2.1), the novices displayed stronger variance in their vertical eye movements compared to the experts and also tended to focus on objects at a greater distance (Chapman & Underwood, 1998).

In basic perceptual psychology research, parameters describing gaze patterns (e.g. the periods of fixation on different objects) are taken as indicators for information processing sequences.[24] At least at the present moment, it is questionable whether gaze patterns could be used to any extent within the framework of a test method to provide a valid assessment of performance: On the one hand, the necessary technical outlay seems very high, despite the availability of promising mobile solutions (e.g. eye-tracking glasses). On the other hand, the recording of gaze patterns only permits evaluation of the chosen visual search strategies; unless it is combined with further methods, it is not possible to determine whether the detected objects are interpreted appropriately with regard to their traffic safety relevance, for example. Nevertheless, the available findings on novice-specific gaze patterns and task-related eye movement measurements could well supply important pointers for the designers of future visualisations in the sense of targeted demand modulation for the development of test items.

*Hazard classification*
Borowsky, Oron-Gilad and Parmet (2009) studied the extent to which differences in a test person's driving experience correlated with differences in their classification of hazard-related traffic situations. Referring to studies conducted by Sagberg and Bjornskau (2006), the authors assumed that at least two factors were crucial for interindividual differences in hazard perception: Firstly, recognition that a situation is hazardous, and secondly the speed with which a person reacts to a perceived hazard. In the opinion of the authors, the focus of research had in the past been placed on the latter aspect, whereas little attention had been paid to the general perception of hazard-related traffic situations.

To answer the question as to whether young, inexperienced drivers (17 to 18 years old with an average 2.7 months of driving experience) classified traffic situations differently to older, experienced drivers (22 to 30 years old with an average 7 years driving experience or 65 to 72 years with an average 38 years of driving experience), the three groups of test persons were initially shown six dynamic traffic situations. The task here was to indicate the detection of a hazard by pressing a button. Subsequently, static images of the same traffic situations were presented; these stills contained the hazard-relevant information cues. The test persons were now asked to group the static images according to freely chosen criteria on the basis of similarities in the hazardous situations. The meaning of "similarity" was deliberately not defined in more specific terms. The image contents, however, already suggested two classification criteria: Classification according to the concrete hazard triggers (e.g. the brake lights of a preceding vehicles) or according to characteristics of the traffic environment (e.g. road in an urban residential area). Different classification patterns were revealed: The young, inexperienced drivers tend to classify the hazards on the basis of the concrete triggers, whereas the experienced drivers also took characteristics of the traffic environment into account. The authors concluded from the study results that it was only with increasing driving experience that not only specific

[24] This indicator function is based on the so-called "eye-mind assumption", according to which fixation (e.g. on an object or a text) always implies conscious processing of the information. While it may be true that – given the contradicting empirical findings (e.g. on change blindness and inattentional blindness, see above) – this hypothesis of a direct connection between eye movement patterns and the associated cognitive processes cannot be considered generally valid, gaze analysis is nevertheless viewed as a reliable methodical approach for the investigation of cognitive processes, especially in the field of traffic psychology (cf. Crundall & Underwood, 2011).

hazard cues, but also potential hazards were perceived; consequently, it was important to address these potential hazards, in particular, in hazard perception tests.

*Risk acceptance*
In a test method serving to assess risk acceptance in road traffic (Vienna Risk-Taking Test – Traffic; Hergovich, Arendasy, Sommer & Bognar, 2007; Arendasy, Hergovich, Sommer & Bognar, 2005), dynamic driving scenarios are presented from the driver's perspective. The test persons are asked to press a button to indicate when they would no longer perform a specified action because it would be too dangerous. The required actions refer primarily to overtaking manoeuvres, speed adaptation and the passing of intersections and junctions, and the video clips are designed such that it becomes increasingly dangerous to perform the action as the scenario develops. The mean latency, i.e. the time from the start of the clip until the button is pressed, is taken as a measure for individual risk acceptance. Each video clip is shown twice within the framework of the test – once for familiarisation with the content, and a second time to record the test person's response.

In the description of their test quality analysis, the authors state that the test assesses a unidimensional latent personality trait which can be interpreted as the "subjectively accepted degree of risk". Through the investigation of correlations between the test results and other variables (e.g. "sensation-seeking"), it was possible to gain evidence for the validity of the method. Despite the fact that the operationalisation described here would evidently offer scope for manipulation in an actual test situation, as candidates could adapt their responses to demonstrate particular safety orientation, the reported findings at least point to the relevance and essential measurability of subjective risk acceptance. Alternative operationalisations are conceivable in that certain scenarios could simply ask whether a certain action can be performed safely at all. The methods of operationalisation chosen in a number of Australian states (see Chapter 2.3.2) provide initial feedback on possible implementations.

*Response latency*
Response latency is a frequently chosen parameter for traffic perception and hazard avoidance performance. Empirical studies which indicate that the speed of reaction to hazard cues increases with driving experience (e.g. Smith, Horswill, Chambers & Wetton, 2009; Scialfa et al., 2011; Scialfa et al., 2012) support the importance attached to an early response to hazard cues in many hazard perception tests. Even so, empirical findings relating to the dependency of response latency on driving experience are by no means unanimous. Sagberg and Bjørnskau (2006), for example, report that the responses to video-based test items relating to hazard perception revealed no significant differences between groups which had held a driving licence for different lengths of time (1.5 months, 9 months, several years).

The various findings on the significance of response latency for valid differentiations of traffic perception competence must also be judged, not least, against the background of the differing conceptualisations of the hazards to be detected in hazard perception tests. According to Borowsky, Shinar and Oron-Gilad (2010), for example, the importance of driving experience is especially relevant for the detection of potential hazards. This also offers an explanation for partly divergent findings relating to response latency: It is possible that the performance differences between experienced and novice drivers are less conspicuous where apparent hazards are to be detected. Instead, corresponding differences are only revealed when the stimulus material contains potentially hazardous situations (Borowsky, Oron-Gilad & Parmet, 2009). On the basis of both the findings reported here and its proven incorporation into actual test procedures (see Chapter 2.3.2), response latency is shown to be a very promising operationalisation approach, as it reflects the speed of the cognitive processing of hazard-relevant traffic information. This speed, it would appear, is not only highly dependent on experience, but can also be acquired with suitable training approaches (see Chapter 3).

*Situation assessment*
Within the framework of a study to test different models by which to explain the higher response latency of novice drivers, Wallis and Horswill (2007) also considered whether experts and novices differed in their subjective assessments of situations. To this end, they investigated three groups of test persons: (1) Novices with a maximum of four years of driving experience, (2) Novices with a maximum of four years of driving experience who had attended an additional video-based training measure, and (3) Experts with more than 10 years of driving experience. The requirement for all three test groups was to react appropriately to the hazard cues contained in dynamic traffic scenarios. In addition, they were asked to use a 20-point scale to rate the probability that traffic conflicts would arise during the further course of the viewed situation. The reactions to hazard cues in the individual groups displayed the expected decrease in response latency with increasing driving experience, but there were no differences in their situation assessments. For the authors, these results support the explanation that novice drivers tend not to react until a hazard cue becomes evident ("response bias model"); experts, on the other hand, already anticipate hazards from the same information cues.

Contrary to the aforementioned findings, Scialfa et al. (2012) discovered differences in the situation assessments of experienced and novice drivers. Their hazard perception test was based on static depictions of traffic situations and required the test persons to rate the hazard risk and the perceived "scene clutter", i.e. the clarity or complexity of the depicted situation, on a 5-point scale (from "not hazardous" to "extremely hazardous" and from "no clutter" to "heavy clutter"). In both areas of the assessment, the ratings of the experienced drivers were on average significantly higher than those of the novice drivers. The differences in situation assessment displayed the expected correlation with response latency, i.e. lower ratings correlated with greater response latency.

Despite the inconsistent findings on the significance of situation assessment for differences in response latency between experienced and novice drivers, both Wallis et al. (2007) and Scialfa et al. (2011) concluded that training concepts for novice drivers should be geared to the early detection of situational information cues and to the development of an elaborated mental model.

*Calibration*
The term "calibration" (Kuiken & Twisk, 2001) refers to the process wherein driving demands are regulated by adapting driving behaviour to the resources available to the individual to cope with any given situation. Key factors for such regulation processes are appropriate self-assessments of driving skills by the driver, correct estimation of the complexity of a driving task, and the selection of behaviour suitable to alter the task demands in an effective manner (DeCraen, 2010).

With the aid of a specially developed test method − the so-called "Adaptation Test" − DeCraen (2010) analysed how calibration changes with driving experience. The Adaptation Test comprises a series of image pairs which differ in only one detail (e.g. a bend to the left on a road outside built-up areas, where the view of the road ahead is either unobstructed or else restricted by bushes). The details concerned contribute to a greater or lesser complexity of the situation in each case. The task for the test person is to specify the driving speed which is considered appropriate in the depicted situation. A response was considered "correct" if the speed selected for the more complex situation was lower than for the less complex situation (an identical speed for both images of a pair and a higher speed in the more complex situation were considered "incorrect"). According to DeCraen, the objective of the Adaptation Test is not merely to assess the detection of hazards; it goes further by asking for behaviour appropriate to the viewed situation. Results obtained from use of the method show that, on average, experienced drivers gave more correct responses than novice drivers. Drivers who overestimated their driving skills in the self-assessment, and likewise those who were identified as unsafe drivers within the framework of an on-road driving assessment, achieved poorer results.

Notwithstanding the very promising nature of the reported study results, there are still certain limitations for the elaboration of a complete model of the construct of calibration using the described method. DeCraen (ibid.) sees one possibility for further development of her method in the use of a driving simulator to visualise the test items. In this way, it would be possible to measure actual driving speed, instead of relying on reported speed behaviour. Furthermore, a driving simulator would enable the assessment of further driving behaviour parameters (e.g. lane-keeping or braking and acceleration behaviour) which are relevant in connection with hazard avoidance and the "calibration" between individual driving skills and situational demands.

### 1.7.4 Principles for the elaboration of test formats and parallel tests

The described methodical approaches and operationalisations for assessment of the various demand components of the construct "traffic perception and hazard avoidance" represent a promising starting point for the development of training offers (see Chapter 3) and test items. When elaborating test formats[25] (and thus the constituent test items and parallel tests), the following principles must be observed:

- *Concrete definition of the test content:* The elaboration of a test concept requires precise definition of the content to be addressed by the test items. Starting points for a corresponding target definition are the construct to be assessed (i.e. its demand facets and structural components to which reference can be made) and possibly existing test methods which can be modified or developed further. It is similarly necessary to take into account the results of research into novice-specific deficits related to the construct.
- *Referencing to learning objectives:* When developing a learning-objective-referenced test, it is necessary to establish references to learning objectives and possibilities already at the stage of task conception. The starting points here are the specifications of the pedagogical and legal control instruments (in other words, the curricular foundations of driving school training and the Learner Driver Training Ordinance), as well as empirical findings on the trainability of the demands to be tackled. The referencing of test items to learning objectives also offers possibilities to transfer content-related and methodical approaches from the construction of a traffic perception test to learner assessments during the course of driving school training and to computer-based media for independent learning.
- *Weighing up of different instruction and answer formats:* Different approaches are conceivable for the visualisation of traffic situations in test item instructions (e.g. static images or dynamic sequences; computer-generated scenarios or video clips recorded in real traffic). These possibilities must be taken into account in

[25] It is not possible to pass a decision in a favour of certain formats within the framework of the present report, as such a decision requires prior empirical testing for the purposes of task and test analysis (as an element of the test development process described in Chapter 2.3.1).

format elaboration, as they each display specific advantages and disadvantages in the context of task and test development. Candidate responses can also be entered in various ways, for example by clicking on a certain area of the screen or indicating an action decision. For the purposes of task design, it seems indispensable to clarify and weigh up the advantages and disadvantages of such operationalisation possibilities very carefully; empirical findings in this area are currently still rather few in numbers (see Chapter 2.3.1).

- *Performance parameters and assessment criteria:* Description of the performance parameters to be assessed (e.g. reaction times, hazard cues detected) is of particular importance for content-related and methodical task conception. For each performance parameter, it is necessary to specify assessment criteria according to which it can be determined whether a task is solved correctly or not; under certain circumstances, this may only become possible in the course of testing of the pre-selected test items.
- *Consideration of evaluation findings:* The scientifically founded use of a certain test format requires that the validity of the format concerned is confirmed by empirical studies. Where evaluations of certain formats already exist, they should be taken into account to enable the selection of promising formats. Further testing of the test items nevertheless remains indispensable.
- *Test economy and technical feasibility:* The elaboration of test items and parallel tests must give due consideration to the aspects of test economy (e.g. the test duration which must be associated with a certain format; the work and expense incurred for elaboration and further development of a pool of test items), in just the same way as to technical prerequisites for the implementation of different task formats. Generally speaking, a computer with monitor and input device can be deemed the basic technical prerequisite. With a view to medium- and longer-term development perspectives, it may be expedient to expand the discussions of task concepts to include also eye-trackers or interactive computer applications with facilities to specify and record steering movements as further technical prerequisites.

The aforementioned principles represent an overarching framework for the description of possible test formats. Lienert and Raatz (1998) point out that as many task concepts as possible should be elaborated at the initial stage, firstly because they will not all be suitable for translation into actual test items, and secondly because a judgement on the usability of individual task concepts can only be reached in the course of implementation testing. A selection of task concepts based on the theoretical and empirical foundations discussed in the present report can be found in the annex to the report.

## 1.8 Integration of different forms of testing in the system of novice driver preparation

One question closely associated with the development of a traffic perception test as a further form of driving licence testing concerns the most meaningful position at which to integrate this new form of testing into the system of novice driver preparation and the process of driving competence acquisition. In this context, it must also be taken into account that tests can only achieve their control function in respect of competence acquisition, if adequate opportunities for learning are available before the test (see Chapter 3).

Thoughts on the placement of a new form of testing are governed firstly by pedagogical-psychological and didactic considerations. At the same time, it seems expedient to examine the circumstances in those countries where a traffic perception test is already implemented in the local system of novice driver preparation. Fig. 13 shows how these countries integrate the traffic perception test into their systems of novice driver preparation alongside the traditional forms of driving licence testing ("theoretical knowledge test", "practical driving test") and opportunities to build up driving practice:

- In Great Britain and the Netherlands, the traffic perception test is taken simultaneously with the theoretical knowledge test at a relatively early stage of the learning process. The novice drivers possess only basic driving experience at the time of the test.
- In the Australian states of Western Australia and Victoria, the traffic perception test marks the transition to the autonomous learning phase. It can be assumed that the novice drivers already possess driving experience at the time of the test, as a period of at least 12 or 23 months of driving practice – and during this period at least 50 or 120 hours of accompanied driving – is a prescribed requirement.
- In the Australian states of Queensland, New South Wales and South Australia, too, candidates must necessarily gain practical driving experience before taking the traffic perception test. The test is here placed in the autonomous learning phase, i.e. after the granting of a driving licence which entitles the holder to drive solo (subject to specific protective measures such as a night driving ban and a zero-alcohol rule). The traffic perception test regulates the awarding of additional driving entitlements, through to the granting of a driving licence without restrictions.

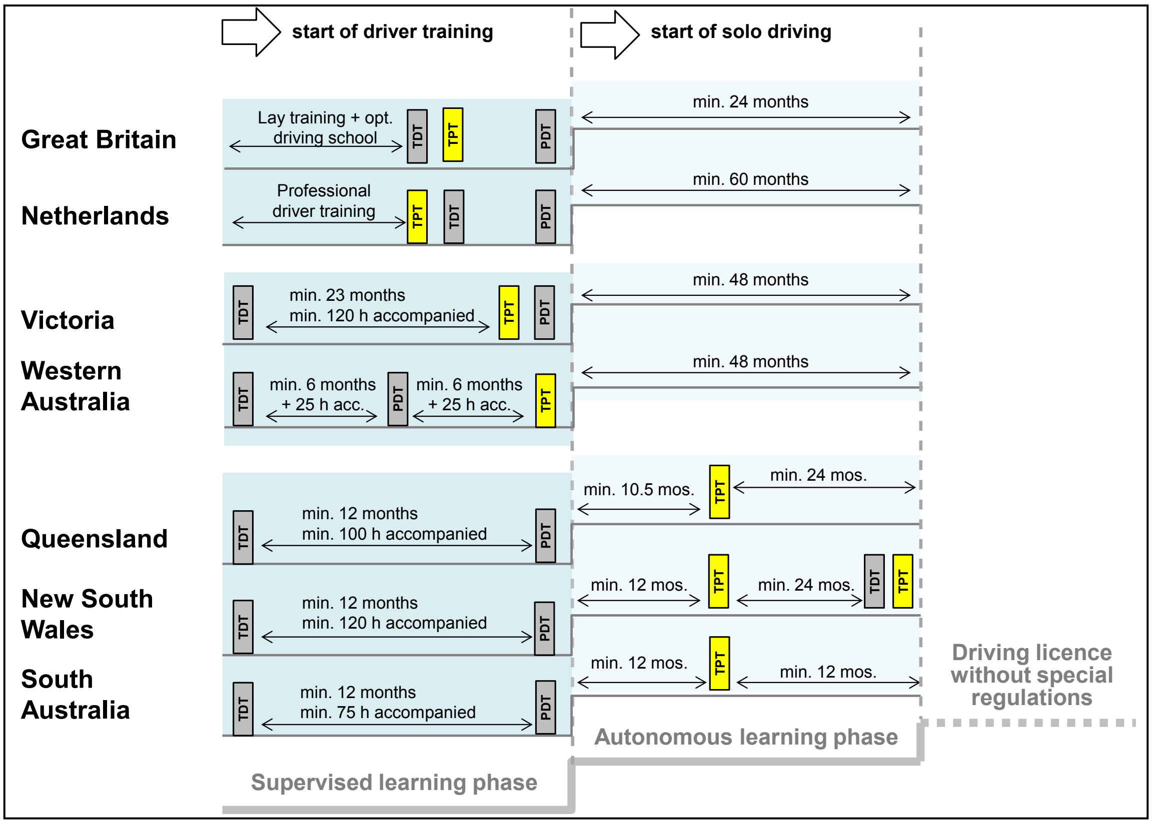



**Fig. 13: Integration of traffic perception tests into systems of novice driver preparation**

Which conclusions can be derived with regard to the placement of a traffic perception test in the system of novice driver preparation in Germany? The TDT and PDT are the forms of testing used in the current Germany system. Despite the fact that certain degrees of freedom exist in respect of preparation for the test and the exact time at which the test is taken, the system architecture already implies a certain sequence of steps in competence acquisition and competence measurement: Both forms of testing are assigned to the so-called “supervised learning phase”, i.e. they must be passed before commencing solo driving. Furthermore, successful completion of the TDT is a prerequisite for admission to the PDT, i.e. the demonstration of successfully acquired traffic-related knowledge always precedes the demonstration of adequate practical driving competence. The PDT, in turn, marks the transition to the so-called “autonomous learning phase”, at least for those novice drivers who do not participate in the “BF17” model (“Accompanied driving from age 17”). The knowledge and skills required to complete the two tests must be acquired by way of formal driving school training.

In the German system of novice driver preparation, therefore, formal driving school training is prescribed for all novice drivers – in contrast to countries with exclusively test-based systems. A traffic perception test would help to expand the acquisition of competence relating to traffic perception and hazard avoidance within this learning process, provided suitable training concepts are made available. These concepts should address not only the theory classes and practical driving instruction, but also independent learning at the computer. The earliest meaningful point at which to plan a traffic perception test, therefore, is at the end of practical driving school training: Firstly, the novice drivers will by then already possess practical driving experience which could be useful for their handling of the test demands; and secondly, the test is then able to realise its control function within the process of driving school training, for example by motivating novice drivers to make use of independent learning opportunities.

In addition, a further developed TDT should in future make a stronger contribution to the acquisition of basic knowledge relating to traffic perception and hazard avoidance. The reasoning behind this demand is that learners are at this point able to acquire the basic knowledge which should preferably be available before their first participation in real motorised traffic – i.e. also before the commencement of practical driving instruction in the driving school – with the aid of virtual traffic scenarios and thus without exposure to real hazards. The successful acquisition of this basic knowledge should thus be verified at the latest upon completion of the theory classes. Accordingly, the TDT incorporates an increasing scope of items which require the application of basic knowledge from the field of hazard avoidance to visualised traffic situations (see Chapter 4).

*Bianca Bredow*

# Concepts for the training of traffic perception and hazard avoidance

## 1.9 Overview

It would seem desirable for learner drivers to be able to acquire competences relating to traffic perception and hazard avoidance in a "protected environment" (e.g. at a computer with the aid of computer-generated visualisations of traffic situations, or in a training vehicle under the supervision of a driving instructor) before being exposed to high accident risks during their first months of solo driving in real traffic. To this end, it is necessary to develop suitable training concepts with relevant content-related learning offers both for the initial phase of independent theory learning and theory classes and for driving simulation training and practical driving instruction.

Which role do training concepts aimed at promoting traffic perception and hazard avoidance currently play in driving school training, and how has their significance changed over the past decades? To be able to answer this question, it is first useful to take a brief look back to the beginnings of driving school training and driving licence testing. It is here immediately conspicuous that legal regulations relating to the obligation for drivers of motor vehicles to hold an official driving licence and to the implementation of a corresponding driving test were already introduced before training possibilities existed for driving licence applicants, and certainly before driving school training became a legal stipulation. In Germany, for example, every driver of a motor vehicle was required to hold an official driving licence from 1900 onwards; it was not until 1904, however, that the first "Chauffeur and Professional Driving School" which was wholly independent of the automobile industry was opened in Aschaffenburg (Sturzbecher, Mönch, Kissig & Marschall, 2009).

Following the introduction of driving licence testing, a prospective driver was required to demonstrate corresponding technical knowledge and acquaintance with the handling of a motor vehicle. In addition, he was expected to display composure, self-control, presence of mind and constant attentiveness (Fournier, 1901). It seems clear that the legal requirement to display "presence of mind and constant attentiveness" – expressed in traffic psychology terms – represented an obligation to observe the traffic environment with due caution while driving and to avoid hazards. Corresponding competence was not a subject of any systematic training, however, and no professionally sound demand and assessment standards existed to enable the determination of such competence within the framework of a driving licence test. In the Berlin police ordinance published in 1901, which was later taken as a model for national regulations on the granting of driving licences, the legislator determined that the driving test must not represent a "mere formal act", and that it must be imperative for the licence applicant to demonstrate that he is "at all times master of his vehicle" (Fack, 2000). Nevertheless, the demands to be satisfied to obtain a driving licence were at this time still minimal. In Berlin, for example, the applicant was only required to drive back and forth in the courtyard of the police headquarters for a few minutes (Brauckmann, Hähnel & Mylius, 2006; Swoboda, 2001). Valid proof of competence relating to traffic perception and hazard avoidance was thus practically excluded.

Even though the number of chauffeur and driving schools grew quickly after 1904, there was still no detailed preparation relating to proper behaviour in road traffic during the course of the driving school training, because a chauffeur was viewed above all in the role of a mechanic; sensibility for the hazards of road traffic was not a priority (Möser, 2004). Furthermore, there was at first some controversy over whether a safety-aware manner of driving could actually be promoted by pedagogical means. Some contemporaries were instead of the opinion that born traits had a greater impact on driving skills: A driver's “natural eye” and “blood-given” awareness of responsibility were deemed more useful than the “good advice” of a driving instructor (Sturzbecher et al., 2009). Very soon, however, the instructions given to driving test examiners regarding realisation of the driving licence test demanded that they pay due attention to the fact that “technical skills and mere knowledge of the police regulations are not sufficient to determine that a person is qualified to drive an automobile on the streets of a city, and must rather be accompanied by further moral traits: Level-headedness, strength of character and sense of responsibility” (Schubert, 2000, p. 242).

The efforts to improve driving licence tests were evidently unable to achieve the desired road safety gains: The number of road accidents increased rapidly in the first decade of the 20th century[26], and the rise in traffic fatalities placed pressure on the authorities to implement corresponding traffic policy measures. Accordingly, the Ordinance on Motor Vehicle Traffic dated 3rd February 1910 specified the demands to be met in the driving test in greater detail and tightened the criteria for passing of the test. In the theoretical part of the test, driving licence applicants were now expected to know how to behave in hazardous situations; in the practical part, the candidate undertook a test drive in real, medium-density traffic, during which he was required to demonstrate the necessary calmness to drive a motor vehicle, alongside a certain minimum presence of mind (Fack, 2000). The first nationally uniform training ordinance was passed in 1921. The teaching and learning content specified in this ordinance was geared exclusively to preparing learner drivers for the driving licence tests. It can thus reasonably be assumed that it is from this point, at the latest, that the first steps towards regular hazard avoidance teaching are to be found in driving school training.

The envisaged transformation of training and test content progressed only slowly, however. It was not until the early 1930s that the focus of driver training gradually shifted away from the technical aspects and towards practical driving-related knowledge and skills (Ostwald, 1931), with training content now paying greater attention to practical driving exercises and the communication of desired traffic behaviour (Mörl, Kleutges & Rompe, 2009). This process lasted several decades and was facilitated, not least, by the increasing convenience in vehicle handling: Innovations such as the “self-starter” (starter motor) and gearbox synchronisation greatly reduced the demands placed on driving licence applicants in terms of technical knowledge (Möser, 2004). Action strategies relating to traffic observation, hazard avoidance and hazard control were thus able to expand their role in training and test content; sophisticated methods for the training and testing of such competences were still lacking, however, as it was still necessary to develop corresponding foundations in traffic and perceptual psychology, as well as traffic-related pedagogical-didactic knowledge and skills, and suitable teaching/learning and test media.

In the early 1970s, the situation slowly began to change; at the same time, the number of persons killed in road traffic in Germany per year reached an all-time high in 1970, with 21,332 fatalities. It is perhaps also for that reason that the topic of “traffic perception and

[26] In 1909, the number of vehicles on the road had approximately doubled compared to 1906, whereas the number of persons killed in road accidents had almost quadrupled over the same period. With a total of 41,727 registered motor vehicles and 194 persons killed in accidents involving motor vehicles, the fatality rate was 46.5 road deaths per 10,000 registered vehicles (Sturzbecher et al., 2009). For comparison, the equivalent rate for the Federal Republic of Germany in 2004 was 1.0 (Statistisches Bundesamt, 2005).

hazard avoidance" was henceforth anchored explicitly in the legal framework governing driving school training. The familiarisation with hazards and possible means to control or avoid hazards were named as essential training objectives (Ausbildungsrichtlinie, 1971). These objectives were only underpinned with suitable training content to a limited extent, however, in spite of the demands expressed by renowned pioneers of the traffic sciences: Gerhard Munsch from TÜV München, for example, saw the key to improved traffic safety in promotion of the "traffic sense" of novice drivers within the framework of their driving school training: The novice drivers should be taught to recognise the signs of a pending hazard at an early stage, and in this way learn to act "in advance of the hazard" ("dynome doctrine"). In an article for the magazine "Der Spiegel" entitled "Driving school: Learn to survive" ("Fahrschule: Lern mal was fürs Überleben", issue dated 26.04.1976), Professor Walter Schneider from Hagen even demanded that the phase of theory classes be extended from six to eight teaching units, in order to permit appropriate treatment of the topic of hazard avoidance.

The stipulations of the current Learner Driver Training Ordinance (Fahrschüler-Ausbildungsordnung, FahrschAusbO) show that the communication of "abilities and skills relating to the perception and control of hazards, including their avoidance and aversion" remains a central objective of driving school training to this day (§ 1 (2) No. 3 FahrschAusbO, 2012). Even so, the legal framework still mentions only a few elements of training content which focus on the attainment of this objective: With regard to theory classes, in particular, there are no systematically structured units devoted to "traffic perception and hazard avoidance", but instead only dispersed points of contact with the topic[27], each of which represents only a small part of the teaching unit concerned.

Over the past years, however, various developments have been initiated to improve the acquisition of competences relating to traffic perception and hazard avoidance. For example, the Federation of Driving Instructor Associations (BVF), the TÜV | DEKRA arge tp 21 working group in Dresden, the Institute for Applied Research on Childhood, Youth and the Family (IFK) at the University of Potsdam and the Institute for Prevention and Road Safety (IPV) in Kremmen have developed two integrated training units to promote traffic perception and hazard avoidance, as well as a traffic perception test with innovative item formats to verify the acquisition of corresponding competences. The training units and traffic perception test are to be tested in four federal states in 2015. Since 2014, furthermore, the learning aid publishers Degener and Heinrich Vogel have offered driving simulators with which learner drivers are in future to learn not least how to observe traffic and avoid hazardous situations. These innovations implement the optimisation potential embodied in the use of new forms of teaching and learning media (Petzoldt, Weiß, Franke, Krems & Bannert, 2011). At the same time, they also reflect the status of research and development in other countries with advanced systems of novice driver preparation.

In the following Chapter 3.2, offers for the training of traffic perception and hazard avoidance which are currently available within the framework of international and national novice driver preparation are described and – insofar as corresponding findings exist – discussed with a view to their learning and safety impact. The chapter is structured

[27] The points of contact within the theory classes include, for example, the teaching/learning component "Hazard perception when using certain types of roads (e.g. tree-lined roads) – Traffic observation, hazard control when changing lanes, tailbacks", which forms Section c of Lesson 4 ("Road traffic system and its use") of the "Framework plan for basic driver training (12 double course units) for all licence classes" in accordance with Annex 1 to § 4 FahrschAusbO (2012).

according to the forms of teaching/learning to which the individual offers are to be assigned: The first section comprises learning offers to support independent theory learning; this is followed by descriptions of offers relating to theory classes and driving simulation training, before a final section which addresses learning offers for the phase of practical driving instruction. The structuring of Chapter 3.2 according to individual forms of teaching/learning is intended to maximise clarity and to facilitate a deeper analysis of the learning offers concerned. For the overall training process, however, the greatest potential for learning gains lies not in isolated consideration of the different forms of teaching/learning, but rather in an integrated approach; corresponding starting points are discussed in Chapter 3.3. In this connection, it must also be taken into account that certain learning offers cannot be assigned unambiguously to one particular form of teaching/learning. Computer-assisted training, for example, is a suitable instrument for both independent theory learning and theory classes, where it can also used to assess learning progress. There is similar no clear-cut distinction between computer-assisted training and (low-cost) simulators.[28]

## 1.10 Learning offers to develop competences relating to traffic perception and hazard avoidance

### 1.10.1 Learning offers for independent theory learning

*Developments in international practice*

Genschow et al. (2013) describe "independent theory learning" as a form of teaching/learning wherein the learning activity is controlled primarily by the learner himself. The learning processes can nevertheless be pre-structured to a varying extent by way of the design of the media used for learning. The media already offer the learner a certain scope of learning content; it is not necessary for a professional instructor to be present. Even so, a professional trainer, e.g. a driving instructor, may still assume an important role, for example through the initiation or indirect guidance of independent theory learning activities (e.g. advice on media selection, specification of preparatory tasks or revision in the formal theory classes). Suitable media are both books and computer-based learning programs.

Subject to the observance of relevant didactic and multimedia design principles, computer-based training programs for independent theory learning, in particular, are deemed to offer considerable potential for the development of competences relating to traffic perception and hazard avoidance (Lonero et al., 1995; Weiß et al, 2009). Firstly, such programs are an inexpensive but nevertheless adequately valid opportunity to practise the application of competences relating to traffic perception and hazard avoidance, without needing to expose the learner to real risks (Wallace et al., 2005; Fisher, 2008; Weiß et al., 2009; Petzoldt et al., 2011). At the same time, they cater to the typically high affinity of

[28] Weiß, Bannert, Petzoldt and Krems (2009) distinguish between low-, medium- and high-cost simulators. Low-cost simulators are understood to be fully functional mock-ups of vehicle operating and display elements which permit traffic scenarios to be presented as large-screen projections. The feedback between the hardware and software is here limited to navigation through the traffic scenario and speed control. In the case of medium-cost simulators, real vehicles are integrated into the simulation; all other characteristics are identical to those of low-cost simulators. High-cost simulators, finally, may take the form of either fully functional mock-ups of vehicle operating and display elements or real vehicles. The difference compared to low- and medium-cost simulators is the dynamic integration of the driver into the simulation by way of a motion platform. In addition to navigation and speed control, the reactions of the vehicle to changes in road position and speed are further levels of feedback between the hardware and software.

the learners – who are most often young adults – to modern media and can thus prove especially motivating (Weiß et al., 2009; Willmes-Lenz, 2010; Schulz-Zander, 2005).

Internationally, numerous training programs for independent theory learning can be found both in the research context and in actual training practice. Most of the programs elaborated for scientific purposes focus on the training of specific components of traffic perception and hazard avoidance. The development of effective scanning strategies here plays a particular role. In this connection, Fisher, Pollatsek and Pradhan (2006) developed the computer-based “Risk Awareness and Perception Training” program (RAPT), which has since been evaluated on several occasions with regard to its effectiveness as a means to support learning (e.g. Fisher, 2008; Vlakveld, 2011).

With the framework of his evaluation of the proposed measure, Fisher (2008) investigated whether novice drivers who had received training based on the training program subsequently displayed better scanning behaviour than their untrained counterparts when driving in real traffic. To this end, 24 novice drivers were assigned randomly to either a control group or an experimental group. While the members of the control group received no specific training, the members of the experimental group were first shown nine sequences of photographs depicting different traffic situations from a driver's perspective, and were asked to indicate the areas of the scene which required particular attention. Subsequently, they were presented schematic bird's-eye views of the same traffic situations, in each case with explanation of the actually relevant areas of each scene (see Fig. 14). Thereafter, the test participants were again required to identify the critical areas in each photograph; the program only proceeded to the next scenario if a correct answer was given. To conclude the training, all nine sequences were shown once more and the areas of potential risk were again to be identified. Overall, the training measure lasted between 30 and 45 minutes.

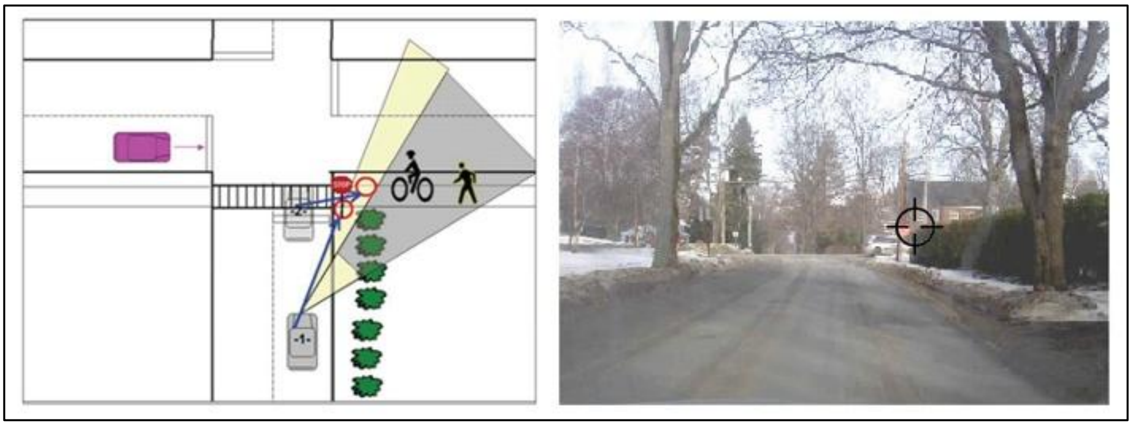


**Fig. 14: Presentations of the “hidden sidewalk” scenario (Fisher, 2008)**

Immediately after the experimental group had completed the training measure, both they and the members of the control group were sent on a test drive along a standardised route in real traffic. During this drive, the participants' eye movements were recorded with the aid of a tracking camera and subsequently assessed in respect of adequacy by three independent observers. The results indicated that the members of the experimental group displayed adequate scanning behaviour significantly more frequently than the members of the control group. One particular difference between the groups was the more frequent fixation on cues to potential hazards by the members of the experimental group than by the members of the control group; this also applied to situations which displayed no similarities to those in the training program (Fisher, 2008).

Fisher did not test whether the described training effects could also be observed after longer intermediate periods. It was nevertheless possible to derive corresponding hints from a variant of the experiment which used a predecessor version of the program and evaluated the training effect on a simulator. In this way, Pollatsek et al. (2006) were able to show that the training effects were independent of whether the simulator test was undertaken immediately after the training or four days later. The authors considered this four-day period of knowledge retention to be sufficient, as this already gave opportunity to realise further training measures (e.g. driving lessons) which could address the same content once more and thus translate the acquired awareness into concrete driving skills (ibid.).

The described findings on use of the RAPT program indicate that computer-based media enable the development of efficient scanning strategies, and that the strategies developed at the computer can also be applied when driving in real traffic. These hypotheses are supported by further scientific studies in which other methods are used to promote effective scanning by novice drivers, for example the commenting of video scenarios (e.g. Chapman et al., 2002; Taylor et al., 2011). With regard to the aforementioned RAPT program, it is especially noteworthy that even instructions of comparatively low physical and functional accuracy are apparently able to achieve improvements in scanning behaviour. Weiß et. al. (2009) view this as a sign that – alongside the media-related implementation of a training program – it is above all the underlying didactic strategy which is of decisive importance for learning success.

At this point, it is useful to consider computer-based training programs which are in actual use to promote traffic perception and hazard avoidance in international practice. Relevant offers for independent theory learning are especially common in those countries in which (1) it is not mandatory to attend theory classes, and (2) appropriate forms of testing ("traffic perception tests", see Chapter 2) are also implemented. This is the case, for example, in Australia ("DriveSmart"; "Shift2ndGear"), New Zealand ("CD Drives") and Great Britain ("The Official DSA Complete Learner Driver Pack", "Driving Test Success"). In the following, the Australian program "DriveSmart" is to be described in more detail; this program has already been the subject of comprehensive evaluation studies.

DriveSmart was developed by the Monash University Accident Research Centre and has been made available free of charge to driving licence applicants in the state of Victoria since 2000. Its aim is to train driving licence applicants in respect of hazard perception, attentional control and the ability to moderate task demands according to their own performance capabilities (Triggs & Regan, 1998). To improve hazard perception skills, for example, the learners watch a series of video clips, in which they are required to indicate when the depicted traffic situation becomes hazardous by clicking with the mouse; subsequently, they must provide further information on the hazards concerned by answering multiple-choice questions. Alternatively, a video clip may be frozen automatically at a relevant point, and the learner must there identify possible hazards and mark them with a mouse click. Feedback is provided after each exercise, and includes both the correct answer and information on the possible consequences of incorrect answers (Petzold et al., 2011). Attentional control is promoted by confronting the learner with the effects of distraction from the task of driving. They are presented two tasks which must be handled simultaneously: While the first task relates to the actual process of driving (e.g. maintaining a constant safe distance to the preceding vehicle in a virtual reality sequence), the second task is completely unrelated (e.g. solving mathematical questions).

The effectiveness of the training was investigated in an experimental study with 103 driving licence applicants (Regan, Triggs & Godley, 2000). Fifty-two participants were assigned to the experimental group and were required to complete five sessions of

DriveSmart training at a computer. The remaining 51 participants formed the control group and occupied themselves not with driver training, but with a flight simulation program; the training contents were selected such that the possibilities for any positive transfer of acquired flying skills to the post-training driving assessment were minimised. The sessions were held at intervals of one week.

The learning effect of the DriveSmart program was assessed with the aid of a driving simulator, in which the driving skills of the participants were measured on three separate occasions: The first measurement took place before the first training session and served firstly to familiarise the participants with operation of the simulator vehicle, and secondly to determine their baseline skill levels by way of an “entrance drive”. A second measurement followed one week after completion of the fifth training session and comprised not only a general practice drive in the simulator, but also special drives which targeted attentional control. It was tested whether the participants were able to share and prioritise their attention appropriately between concurrent tasks, in this case by adapting their speed to a succession of speed limit signs while at the same time solving arithmetic tasks. Furthermore, each participant now completed several special drives which served to determine their abilities in respect of hazard perception, assessment and avoidance. These drives included both new traffic situations (assessment of the “far transfer” effect) and situations which had been encountered during the DriveSmart training and were thus familiar to the participants in the experimental group (assessment of the “near transfer” effect); in some cases, the tasks concerned were again to be performed parallel to the solving of arithmetic tasks. The third measurement, finally, was scheduled four weeks after the completion of the training measure; like the previous measurement, this comprised a practice drive and tasks relating to attentional control and hazard perception. Furthermore, each participant completed an “exit drive”, which was identical to the entrance drive during the first study session and served to permit a before-and-after comparison of the participant's performance (Regan et al., 2000).

The results of these studies showed that the members of the experimental group displayed greater risk awareness during the simulator drives and were better able to avoid arising hazards than the members of the control group. No differences were found between the handling of traffic situations which were known from the training program (near transfer) and the handling of new situations (far transfer). The members of the experimental group also performed better than the control group with regard to attentional control, and the differences between the two groups were for the most part still present at the time of the third measurement four weeks after completion of the training (Regan et al., 2000). However, the study does not permit conclusions to be drawn as to the extent to which skills acquired through the computer program and demonstrated in the simulator are also transferable to driving in real traffic.

*Developments in Germany*

In Germany, computer-based offers relating to the training of traffic perception and hazard avoidance have to date been developed primarily for use within the framework of scientific studies. Petzoldt et al. (2011), for example, devised a training program in accordance with general teaching and learning theory principles to investigate the potential offered by computer-based learning media for the promotion of traffic perception and hazard avoidance. The program elaborated for this purpose was built around implementations of the VICOM software tool developed by the TÜV | DEKRA arge tp 21 working group and comprised a 10-minute introductory presentation, which served to motivate the participants and communicated necessary skill-specific prior knowledge (instructional element), and two 45-minute training units (active elements). Each training unit, in turn, comprised 13 dynamic scenarios, which were presented together with a variety of multiple-choice questions and marking tasks. Based on the different levels of situation awareness (Endsley, 1995; see Chapter 2), three types of question were used:

(1) Questions relating to object perception (e.g. “Where are any other vehicles relative to your own position?”),

(2) Questions relating to evaluation and projection (e.g. “Mark the areas to which you should pay particular attention!”, see Fig. 15), and

(3) Questions relating to behaviour and understanding (e.g. “Which reaction is appropriate? Why?”).


**Fig. 15: Example of question type 2 from the learning program (Petzold et al., 2011)**

To assess the effectiveness of the learning program, Petzoldt et al. (2011) recruited 58 learner drivers and assigned them to three different groups: The members of the first group completed the computer-based training program in addition to their driving school training; they worked with dynamic scenarios and received differentiated, error-specific feedback. The members of the second group were given a printed learning book to supplement their driving school training. In terms of content, the exercises in the learning book corresponded to those of the computer-based training program; the driving situations were presented as static images, however, and the learner was offered only a general indication of the correct solution rather than error-specific feedback. The members of the third group, finally, received no additional training alongside their driving school training. The chosen study design was intended to permit assessment of both the effectiveness of the learning software and possible influences of the learning medium.

Two days after the intervention, all learner drivers completed a test drive in a simulator. One of the purposes of this drive was to analyse the glance behaviour of the learner drivers.[29] At the same time, it was determined whether the learner drivers displayed appropriate driving behaviour during the simulated drive.[30] Evaluation of the results showed that those learner drivers who had trained with the aid of the computer-based software dis-

[29] Glance behaviour was recorded with the aid of an eye-tracking camera. The variables used for operationalisation included the places, numbers and durations of fixations, as well as the latencies and contingencies between different points of glance focus.

[30] The authors considered “appropriate” driving behaviour to mean behaviour which “serves to avoid accidents and furthermore, insofar as the concrete traffic constellation permits, does not hinder driving flow” (Petzold et al., 2011, p. 29). No more specific definition was given, because different forms of driving behaviour were deemed appropriate in different scenarios.

played significantly better glance behaviour than the members of both other groups. Moreover, their glance behaviour was similar to that of experienced drivers. At the level of driving behaviour, however, there were only a few test scenarios which produced significant improvements compared to the other groups (Petzoldt et al., 2011). According to the authors, this was in particular due to the high number of degrees of freedom in the way the driving scenarios could be handled. Furthermore, this result could be attributed to the design of the multimedia learning offer, which forewent the integration of dynamic interactions (e.g. pressing of a button or click with the mouse upon appearance of a hazard cue) as reactions in the training scenarios.

No significant differences were found between the learner drivers who had used the printed learning book and those who had received no additional training, neither in terms of glance behaviour, nor with regard to overall driving behaviour. The authors thus conclude that the learner drivers who had used the printed learning book, despite their comprehensive knowledge of particular traffic situations and the correspondingly required behaviour, "are not (yet) able to apply this knowledge productively" (ibid., p. 36). Conventional learning media – in contrast to computer-based media which have been designed according to pedagogical and didactic principles, and as such offer new possibilities for presentation and interaction with the learning content – were thus seen to offer little potential for the acquisition of practice-relevant traffic perception and hazard avoidance skills.

Besides Petzoldt et al. (2011), Hilz, Malone and Brünken (2014) are also currently working on the development of computer-based training programs for use in driving school training. Following the four-component instructional design model (van Merriënboer, Clark & de Crook, 2002) and building upon three empirical studies, they elaborated a training software which was to be used specifically to train hazard perception. On the basis of accident analyses, driving task concepts, findings from teaching/learning research and the differences between novice and experienced drivers, they first used the VICOM software provided by the TÜV | DEKRA arge tp 21 working group to develop different task formats (multiple-choice questions, timed-reaction tasks, marking of critical features). For the first study, they recruited 120 participants to test whether and, if so, to what extent the elaborated items were suitable to distinguish the performances of novices and experienced drivers. The second study continued by determining which additional information needed to be made available to the learners in which form in order to raise the effectiveness of the training. The third study, finally, combined the outcomes of the two preceding studies: Tasks which had yielded the greatest differences between novice and experienced drivers in the first study were expanded with additional information in accordance with the findings of the second study and then offered to the 60 learner drivers of the experimental group alongside their driving school training. Another 60 learner drivers formed the control group; they received no further training in addition to their driving school training. The authors presume that, after completing the training, the learner drivers in the experimental group will display both more appropriate glance behaviour (e.g. numbers and durations of fixations) and better driving behaviour (fewer errors in the handling of traffic situations) than the learner drivers in the control group. Final empirical results, however, were not yet available at the time of elaboration of the present report.

It can be taken from the above that, particularly in recent years, the potential of computer-based learning offers as a means to promote traffic perception and hazard avoidance within the framework of independent theory learning has been shifted into the spotlight of scientific discussion in Germany. The extent to which such programs have already entered German driving school practice, however, remains uncertain. Answers to this question are sought in the following.

In Germany, learner drivers prepare for the theoretical driving test primarily by attending mandatory theory classes. Independent theory learning can be used to supplement the theory classes or for preparation and revision of the discussed content. In contrast to other countries (above all Finland and various US states), however, the integration of independent theory learning into the German system of driving school training has to date remained essentially unsystematic: The corresponding learning offers are generally voluntary and their use depends on the learning preferences and personal commitment of the individual learner driver (Weiß et al., 2009).

For a long time, self-learning programs which were geared solely to familiarisation with the official test questions, and thus the acquisition of fundamental factual knowledge, predominated (Prücher, 2006). These programs were by all means suitable to support individual preparation for the theoretical driving test, but were unable to contribute to the development of practical action competences, e.g. competences relating to traffic perception and hazard avoidance. There was no systematic construction of computer-based self-learning programs designed specifically to develop such competences (Weiß et al., 2009). In the current self-learning programs, too, the main focus is placed on the communication of factual knowledge. In the meantime, however, the first supplementary offers aimed at improving traffic perception and hazard avoidance are also to be found in the portfolios of the major German teaching material publishers. With its collection of visual perception exercises and so-called "practice packages", the publishing house Verlag Heinrich Vogel, in particular, offers materials to support not only the acquisition of factual knowledge, but also the promotion of practice-relevant competences such as correct observation of the traffic environment, the identification and assessment of hazards, and the determination of appropriate actions for hazard avoidance.

### 1.10.2 Learning offers for theory classes

*Developments in international practice*

Genschow et al. (2013) describe theory classes as a form of teaching/learning in which traffic-related knowledge is conveyed to learner drivers by a professional instructor. Compared to independent theory learning, theory classes are credited with much greater potential to convey not only skills relating to traffic perception and hazard avoidance, but at the same time also corresponding, safety-enhancing attitudes, especially when discursive and special attitude-building learning methods and innovative media are employed.

The extent to which the topic of "Traffic perception and hazard avoidance" is anchored in the content addressed by theory classes in other countries was studied within the framework of the BASt project "Approaches to the optimisation of driver education in Germany" by way of a comparative analysis of 14 progressive driver training curricula implemented in international practice (Bredow & Sturzbecher, in press). The analysis results suggest that the development of competences relating to traffic perception and hazard avoidance is a central topic and, as such, the subject of separate course units in many driver training curricula (e.g. Dubai, Ireland, Ontario, Québec). In Québec, for example, driving licence applicants must complete separate theory modules dealing with novice-typical risk behaviour (e.g. speed regulation). In Ontario, a similar course module "Perception and risk management" is devoted to observation techniques, individual risk factors, accident avoidance strategies, and strategies for hazard perception and hazard avoidance. The Irish "Steer Clear" programme (Irish Drivers Education Association, 2006), finally, assigns more than half of its (theoretical) training to hazard-related topics: Driving licence applicants must here complete six training modules on subjects such as "Attention", "Hazard perception", "Risk evaluation", "Motor skills", "Personal risk acceptance" and "Mastering risk situations".

*Developments in Germany*

In Germany, in contrast to other countries (see above), there are to date no course units geared specifically to the promotion of competences relating to traffic perception and hazard avoidance. Petzoldt et al. (2011) criticise the fact that the treatment of hazard avoidance in theory classes has so far remained limited to the communication of fundamental declarative knowledge. To overcome this safety-relevant deficit, the Federation of Driving Instructor Associations (BVF) – in cooperation with the Institute for Applied Research on Childhood, Youth and the Family (IFK) at the University of Potsdam – elaborated a 90-minute training unit "Traffic perception and hazard avoidance in road traffic" within the framework of the BASt project "Approaches to the optimisation of driver education in Germany" (Bredow & Sturzbecher, in press). This training unit serves to convey general foundations to the learner drivers, to help them better recognise, assess and avoid hazards. The focus is placed on five individual topics, which are handled by way of discursive teaching/learning forms and interactive media:

- Deficits in traffic perception as a cause of accidents,
- Hazards in road traffic,
- Strategies for effective and efficient traffic observation,
- Strategies for hazard assessment, and
- Strategies for hazard avoidance.

Building upon the aforementioned basic training unit on general hazard avoidance, the Institute for Prevention and Road Safety (IPV), the TÜV | DEKRA arge tp 21 working group and the Federation of Driving Instructor Associations (BVF) then developed a supplementary training unit "Risks for young novice drivers and regional accident black spots". In this training unit – in line with the road safety project "Regio-Protect 21" (Bredow, 2014), as implemented in the federal state of Brandenburg since 2008 – regional accident prevention stands in the foreground. During the theory classes, roads passing local black spots at which particularly high numbers of young drivers are involved in accidents are first "driven" virtually by viewing video clips, and then analysed with regard to possible hazards and potential causes of accidents. To this end, "accident reports" are made available as a source of authentic information on the persons involved and the surrounding circumstances (e.g. weather conditions, time of day) of novice driver accidents on the roads concerned. On this basis, the driving licence applicants are then expected to apply and expand the acquired knowledge and skills in the course of practical driving instruction. The experience gained in this way is subsequently reviewed together with the driving instructor; in particular, common points are sought in the knowledge derived from the virtual and real drives. The route examples are intended to train the learner driver's general traffic perception and enable the selection of appropriate hazard avoidance strategies in any given situation. Evidence for the motivational power and learning effectiveness of this training approach was provided by an evaluation study with 42 learner drivers (Bredow, 2014).

In the meantime, descriptions and foundations have been elaborated in accordance with traffic psychology principles for both training units, together with tabular course plans and teaching/learning media (e.g. teaching presentations, computer-based exercises, worksheets). With regard to the second training unit, furthermore, didactically formulated training instructions have been developed as a basis for corresponding elements of practical driving instruction on the black spot roads. At the same time, demonstration films have been produced (see Fig. 16) to illustrate the different didactic phases of the training units and to present possibilities for practical realisation of the training concerned.

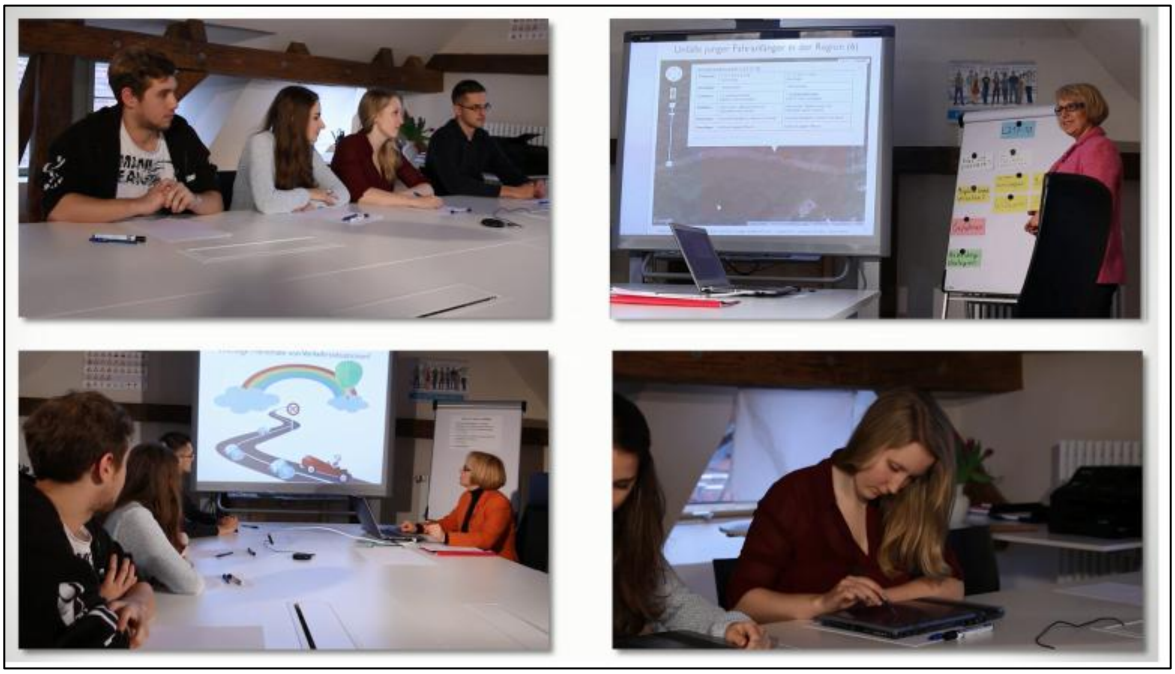


**Fig. 16: Scenes from the demonstration films on the training units “Traffic perception and hazard avoidance in road traffic” and “Risks for young novice drivers and regional accident black spots”**

On the basis of the aforementioned materials, both training units are to be tested in a transregional study in 2015, in order to gain knowledge on their learning effectiveness. The study is to be conducted by the Federation of Driving Instructor Associations (BVF), the TÜV | DEKRA arge tp 21 working group, the Institute for Applied Research on Childhood, Youth and the Family (IFK) at the University of Potsdam and the Institute for Prevention and Road Safety (IPV). The expected learning gains are to be verified by way of (1) a knowledge test, (2) a traffic perception test using innovative question formats to test competences relating to traffic perception and hazard avoidance, and (3) a drive in real traffic to enable driving behaviour observation and assessment using the electronic test report of the optimised practical driving test. In this way, the planned project establishes links between driving school training and the driving licence test by applying common learning standards, alongside the demand and assessment standards defined by the catalogue of driving tasks (see Chapter 6).

## 1.10.3Learning offers for driving simulation training

*Developments in international practice*

Simulators are a means to communicate learning content and to monitor learning progress. The content addressed in a particular case can also be selected according to the current level of ability of the learner driver (van Emmerik, 2004). With these possibilities, and given the vivid and realistic nature of the stimulus material, simulators are at the same time especially motivating for the learner (Weiß et al., 2009). Furthermore, they permit the use of visual instructions and feedback (e.g. indication of an optimum driving line on the screen) and objective recordings of learning progress (Genschow et al., 2013). Finally, simulators offer possibilities to confront learner drivers with potentially hazardous situations and the consequences of an incorrect behaviour choice without exposing them or other road users to real risks (Weiß et al., 2009). The extent to which this is able to promote competences relating to traffic perception and hazard avoidance on the part of learner drivers has been investigated in various scientific studies; two such studies are to be viewed in more detail in the following.

Specifically with regard to glance behaviour, Carpentier et al. (2012) developed a simulator training unit and tested learning effectiveness by way of a experimental group/control group comparison with 28 study participants. The members of the experimental group completed the elaborated training unit, which began with a simulated drive through a scenario of ten potentially hazardous situations; the actual hazards, however, did not (yet) materialise during this drive. At the end of the simulated drive, the participants were asked to indicate any moments during the drive at which they expected a hazard to arise. This question served to stimulate self-reflection for the case of any later accident (or near-miss) and to avert the attribution of an accident to other factors, e.g. other road users (Vlakveld, 2011).

After this discussion, the members of the experimental group completed a second simulated drive. This time, all the potential hazards actually materialised. Each time the participant encountered a hazardous situation, irrespective of whether it was handled more or less successfully, he was asked to stop the vehicle at the side of the road and was shown images depicting the hazardous situation concerned as a bird's eye view and from the driver's perspective. On the basis of these images, the driving instructor then provided verbal and written explanations of the elements of the situation which should have been observed and how the driver could have avoided the hazard. Subsequently, the participant returned to the simulator and was required to handle the hazardous situation once more in order to apply and consolidate the acquired knowledge. Only then was the drive continued to the next hazardous situation, where the aforementioned procedure was repeated.

The members of the control group also completed two simulated drives. The difference compared to the experimental group was that the members of the control group were not asked to reflect on possible hazards after the first drive. Furthermore, no actual hazards were encountered during the second drive; as a result, the members of the control group also received no explanations relating to hazardous situations. Instead, they simply answered two general traffic questions at the relevant points of the simulated drive.

The learning effectiveness of the training was evaluated by analysing the results of two simulator sessions, one immediately after completion of the training and the second two to four weeks later. In both cases, it was determined firstly whether the participants were capable of handling hazardous situations (operationalised by way of the number of virtual accidents caused). In addition, the extent to which they displayed adequate glance behaviour was assessed. Glance behaviour was here operationalised by way of the time at which a hazard was detected, the correct consideration of so-called “visual search points” (e.g. traffic signs, children at the other side of the road) and use of the vehicle mirrors. The test results show that the members of the experimental group scanned the relevant “visual search points” more frequently than their counterparts in the control group, and that they also used the rear mirrors more often. These differences can be interpreted as signs for the acquisition of systematic glance behaviour. On the other hand, there were no significant differences with regard to the number of accidents caused and the time at which hazards were first detected. The authors attribute this result to the small number of study participants (Carpentier et al., 2012).

In view of the variety of existing simulator types (e.g. low-, medium- and high-cost simulators, see above), it can be asked whether the so-called “fidelity”, i.e. the degree of realism achieved with a simulator (near-realistic features and operating controls comparable to those of a real vehicle) is decisive for the learning effectiveness of simulator-based training measures. Allen et al. (2007) considered this question within the framework of a longitudinal study with a total of 554 learner drivers. To this end, two high school classes trained with a low-fidelity simulator comprising a monitor, a steering wheel and pedals. Further study participants, who were recruited through the local vehicle and driver licensing authority, were trained in the laboratory on two simulators with greater degrees of realism. One of these simulators comprised three monitors, a steering wheel and pedals; the second comprised a video projection screen and a static vehicle mock-up.

During the training program, the study participants acquired knowledge relating to the subjects “Rules of the road”, “Lane changing”, “Turning and use of turn indicators”, “Situation awareness and defensive driving” and “Hazard recognition” (Allen et al., 2007). The acquired knowledge was subsequently consolidated over the course of six practice drives, each of which lasted between 12 and 15 minutes. All participants received corresponding feedback on their performance (e.g. speed limit violations, failure to observe traffic lights and signs, use of the brakes). If a participant failed to meet predefined performance criteria on the first run-through, three further attempts were granted. If the criteria were still not met after the additional attempts, the participant was admonished to drive more carefully in the future and took no further part in the study.

To evaluate the training effect, the accident data of the study participants were collected for a period of two years after they obtained a driving licence. It was revealed that a correlation existed between the assignment to a particular training group and the accident rates: The accident rate for those participants who had been trained using a single monitor with corresponding control elements (approx. 7.8%) was twice as high as that of the group which had trained in the vehicle mock-up (approx. 3.5%). All the study participants displayed a lower accident rate than novice drivers who had not completed any special training (approx. 9%; Allen et al., 2007).[31]

Despite the statistical significance of the findings and the considerable size of the subject population, the study results are still unsuitable to prove a positive overall effect of simulator training, as the authors used comparative data for novice drivers who lived in different regions to the training groups. The accident rates for novice drivers without special training thus referred to a non-comparable traffic environment. Furthermore, the assignment of the study participants to a particular set of training conditions was not random. For practicability reasons, school classes were recruited for the training group which used the single-monitor simulator, and it is thus not possible to exclude confounder effects in the study results, for example influences attributable to education background or socio-economic status (Ewert & Steiner, 2013). On the basis of an analysis of various studies, De Winter et al. (undated) argue, furthermore, that the success of a simulator-based training measure is dependent not so much on the realism of the simulation, but rather on its didactic integration into the training process.

What importance is attached to driving simulation training in international novice driver preparation? Answers to this question are provided by a study conducted by Genschow et al. (2013), in which the systems of novice driver preparation in 44 countries were analysed with regard to the forms of teaching and learning used. In none of the countries considered is driving simulation training a mandatory component of driver training. There are only a few countries, such as the Netherlands, Finland and the Czech Republic, in which driving simulator sessions – subject to observance of certain technical feature standards – may be acknowledged in lieu of driving lessons in real traffic. The driving licence applicant is himself able to decide whether to transfer part of the training to a

[31] The percentages represent the number of accidents per 100 persons per year.

simulator or else to complete the full course of practical driving instruction in real traffic. Driving simulation training is usually used at the beginning of practical driving instruction and there serves primarily to familiarise the learner driver with the traffic environment, the vehicle controls and the execution of elementary driving manoeuvres. According to estimates provided by experts, however, only a small proportion of learners actually choose to use a driving simulator (e.g. 15% in the Netherlands; ibid.).

Weiß et al. (2009) view the embedding of driving simulators into “integrative learning systems” as the decisive criterion for the success of their use in driver training. In other words, learning success is dependent on the displayed driving performance being analysed together with the driving instructor, integrated with other forms of teaching/learning (theory classes, practical driving instruction) and taken as a basis for the planning of further training elements. The extent to which this demand is met in practice varies from country to country.

In the Czech Republic, learner drivers are only permitted to train with a simulator before commencing practical driving instruction, i.e. before driving in real traffic. It is possible to complete up to 10 of the prescribed 28 hours of practical driving instruction on a driving simulator. Driving simulation training generally places skills relating to vehicle operation, traffic observation, the performance of specific driving manoeuvres, compliance with traffic signs and rules, and the mastering of risk situations in the foreground, but no cross-references are established to other forms of teaching/learning.

In the Netherlands, too, driving simulation is usually used at the beginning of driver training, i.e. before driving in real traffic. Here, the learner driver's driving behaviour data (e.g. visual attention) are recorded and compared with results of driving behaviour analyses based on the data of around 10,000 other novice drivers. On the basis of this comparison, the learner driver receives detailed feedback on his driving performance. The same feedback is also communicated to the driving instructor, who is then able to take the simulator results into account when planning the further course of the training (Genschow et al., 2013).

*Developments in Germany*

In Germany, practical driving instruction takes place primarily in real traffic (Genschow et al., 2013). Since 2014, learning aid publishers Degener and Heinrich Vogel, in particular, have offered driving simulators which can be used in preparation for driving in real traffic. The use of driving simulators is not governed in any way by legal specifications; this means that the completion of simulator training does not lead to reduction of the required minimum number of on-the-road driving lessons.

With regard to the teaching/learning content, both driving simulation variants focus on aspects of basic vehicle operation (e.g. familiarisation with the vehicle controls, correct use of the gear lever and steering wheel), the training of turning procedures, and behaviour relating to right of way. As the learner driver works through the individual modules, his driving behaviour – and in part also his glance behaviour – is recorded and assessed (e.g. mirror use, shoulder checks, speed adaptation). On this basis, feedback is provided with regard to driving errors; it is intended that the driving instructor takes this feedback into account when planning the practical driving instruction. Concrete recommendations on how the feedback is to be taken into account, and how driving simulation training can be integrated with other forms of teaching/learning would be desirable.

Which conclusions can be drawn with regard to the future use of driving simulators in novice driver preparation? Despite the diverse potential of driving simulators, it would seem that their practical significance for novice driver preparation in general and for the training of competences relating to traffic perception and hazard avoidance in particular is (still) limited. This can be attributed above all to the relatively high investment costs for the nationwide introduction of training simulators. Technically less sophisticated PC-based driving simulators, on the other hand, already today offer widespread possibilities for the training of those partial competences (e.g. traffic perception) which are relevant for the mastering of driving tasks. As long as no findings are presented to show that true simulators enable a greater learning effect than PC-based applications, or else that they achieve comparable effects in a shorter time, it can be presumed that – already from the perspective of efficiency – simulation-based learning software will continue to prevail in the future (Weiß et al., 2009).[32] Furthermore, the existing simulators fall short of exploiting their full potential with regard to hazard avoidance (e.g. learners are still too seldom confronted with situations depicting the development of potential hazards which, although they may occur relatively infrequently in real traffic, must be associated with a high risk of accident if they are not detected and avoided in good time). Finally, the stress symptoms such as nausea, disorientation, headaches or dizziness (so-called “simulator sickness”), which are to be expected among a certain proportion of the learner drivers during simulator use, also hinder the systematic introduction of driving simulators in driving school training (Hoffmann & Buld, 2006; Weiß et al., 2009).[33]

### 1.10.4Learning offers for practical driving instruction

Genschow et al. (2013) describe “practical driving instruction” as a teaching and learning form which comprises mainly in-vehicle teaching/learning situations, in which application-related skills (e.g. vehicle operation or vehicle handling in particular traffic situations) are conveyed by an instructor (e.g. a driving instructor or a lay trainer). The practical driving instruction is given predominantly on public roads, but in part also on separate practice grounds. One core objective of practical driving instruction has always been the systematic acquisition of competences relating to traffic perception and hazard avoidance. This competence acquisition is hampered, however, by the fact that the learning opportunities in the real traffic context are dependent on the randomly encountered circumstances (Petzoldt et al., 2011).

One possibility which has been in use for some time to promote competences relating to traffic perception and hazard avoidance in real traffic is the method of “commentary driving” (sometimes also “commented driving”). Seidl and Hacker (1991) understand this to refer to a teaching/learning method which requires the learner driver – during a practice drive – to verbalise perceptions relating to information from the traffic environment which is relevant for his driving behaviour in abbreviated form. In part, the driver should similarly express his expectations regarding the possible behaviour of other road users and the then necessary reaction on his own part (Gürten, Neumeier & Wiegand, 1987). Such commentary is for the driving instructor an efficient means to assess the driver's learning progress and to identify any deficits in respect of his information acquisition, hazard anticipation or hazard assessment, as a basis for corresponding evaluation feedback (Russell, 2003). Furthermore, the commentary task is expected to counteract distraction effects on the part of the learner driver (Marek & Sten, 1972).

---

[32] Regarding driving licence testing, it is to be added from the methodical perspective, that the apparently greater realism which can be achieved with a simulator compared to PC-based test approaches would only prove a benefit if leads to more valid test decisions. Empirical evidence for such added value has not been furnished to date.

[33] Simulator sickness is attributable to discrepancies between visual perceptions and the information received from the sense of balance.

Crundall, Andrews, van Loon and Chapman (2010) examined the effectiveness of "commentary driving" within the framework of an empirical study based on the experimental group/control group design. All the study participants first completed a drive in the simulator, during which they encountered three hazards from each of the categories "Behavioural prediction", "Environmental prediction" and "Dividing and focusing attention".[34] At the same time, their driving behaviour – operationalised by way of their average speed, speed adaptation when confronted with a hazard, and their use of the brake and accelerator pedals – was recorded. Following the simulator drive, the members of the experimental group met in the classroom for a lesson in which they learned the basics of commentary driving. Subsequently, a two-hour training session took place in real traffic. The participants were accompanied by specially trained driving instructors, who first demonstrated the method of commentary driving, and then provided feedback on the commentary given by the participants. The members of the control group received no training after the simulator drive.

Approximately two weeks after completion of the training, the learning growth of all study participants was tested by repeating the simulator drive with which their initial competence had been measured. It was here found that the members of the experimental group caused fewer accidents during the second drive and also adapted their speed to an arising hazardous situation earlier than the members of the control group (Crundall et al., 2010). The same did not apply for hazards of the category "Environmental prediction", however. The authors interpret the results as evidence that the members of the experimental group had been moved to a deeper analytical engagement with the conditions of road traffic through the previous verbalisation of arising traffic situations. This enables them to perceive hazards earlier and to initiate appropriate behaviour to avoid hazards in good time. The extent to which this competence is preserved in the longer term, on the other hand, was not discussed by Crundall et al. (2010).

The effectiveness of commentary driving for the promotion of competences relating to traffic perception and hazard avoidance is corroborated by further studies (Deery, 1999; Horswill & McKenna, 2004; Seidl, 1990; Wetton et al., 2013; Williamson, 2008; Isler, Starkey & Williamson, 2009). It remains to be added, however, that the effects of commentary driving were not determined in a "live commentary" situation in the aforementioned studies. It is without doubt that commentary driving – when to be handled parallel to the immediate driving task – places additional demands on the learner driver, such that the required verbalisation of the traffic situation could lead to interference effects and increased workload (Helman, 2008). Current findings by Young, Chapman and Crundall (2014) support this presumption: The authors determined the impact of live commentary concurrent to the necessity to handle a hazardous situation. It was shown that the commentary could be associated with extension of the time which passed before the driver reacted to a hazard.

Even so, commentary driving has established itself in international training practice: It is used, for example, in Belgium, Germany, Great Britain, Iceland, the Netherlands, Austria

[34] Hazard situations of the category "Behavioural prediction" contained visible hazard subjects whose behaviour had to be anticipated (e.g. playing children, who could suddenly run onto the road). "Environmental prediction" meant that only the context of a hazard was depicted, while the hazard subject itself remained invisible (e.g. a bus at a bus stop). In the category "Dividing and focusing attention", several potential hazards were shown at the same time, one or several of which could develop into an acute hazard. The task for the participant was to divide his attention optimally, in order to be able to recognise the development of an acute hazard in good time (Crundall et al., 2010).

and a number of Australian states. In each case, either the instructor (e.g. Netherlands, Victoria) or the learner driver (e.g. Germany, New South Wales, Victoria) comments orally on his observations and the hazards recognised in the course of a drive (Genschow et al., 2013; Bredow & Sturzbecher, in press). In Victoria, commentary driving has been developed into a three-stage process with alternating roles: First, the instructor demonstrates the execution of a given driving task, and while doing so describes his observations and actions out loud. Subsequently, the learner completes the same driving task, during which the instructor again comments on his observations and actions. Once the learner has mastered the driving task under normal traffic conditions, finally, he must himself try to give a spoken description of his observations and behaviour while performing the driving task.

The possibilities for the use of commentary driving are not limited to the element of practical driving instruction in real traffic. This teaching/learning method can also be used in theory classes or independent theory learning, for example where hazardous routes are "driven virtually" before the learner driver tackles them in real traffic. In this way, the benefits of the method – and in particular its potential for the acquisition of competences relating to traffic perception and hazard avoidance – can be utilised without the risk that the aforementioned greater workload placed on the learner driver could lead to actual endangerment in real traffic.

## 1.11 Outlook

In the meantime, there is well founded scientific evidence that learner drivers can be prepared for demands relating to traffic perception and hazard avoidance with the aid of pedagogically sophisticated training materials, and that already before they gather practical driving experience in real traffic. Accordingly, various concepts are currently being developed in Germany, with the objective of integrating the aspect of hazard avoidance more strongly into independent theory learning, theory classes and driving simulation training. These concepts are intended to supplement the hazard avoidance components of practical driving instruction and are to be developed further to this end, also taking into account the experience gained in international practice to date.

For the further development of training concepts, it is especially important to consider research findings in the field of teaching and learning theory. Analyses of driving competence acquisition, for example, suggest that training concepts with distinct segmentation into instructional and active learning elements are detrimental for learning success (Grattenthaler et al., 2009). Referring specifically to the area of traffic perception and hazard avoidance, Meir et al. (2010) conducted studies to determine the form of learning which is best able to promote the perception and scanning skills of learner drivers. To this end, inexperienced drivers completed either exclusively instructional training, active training without feedback or a combination of both forms. At the end of the course of training, all participants completed a traffic perception test which assessed their visual search patterns and their competence in respect of hazard anticipation. Whilst the participants who received solely instructional training showed no learning effects, the members of the active learning group displayed eye movements and anticipation skills similar to those of experienced drivers. The best training results, however, were recorded by the participants from the "hybrid training" group – especially when a certain sequence of training elements was observed.[35] The authors concluded that effective training concepts to promote traffic perception must include both instructional elements for knowledge acquisition and active elements to enable practical knowledge application. The better results of the “hybrid training” group compared to the participants who had received only active training were attributed to the fact that, thanks to the instructional component in their training, the former were more successful in associating hazard cues with an actual hazard (Meir et al., 2010).

From the perspective of teaching and learning theory, it is furthermore deemed expedient to view the learning offers for independent theory learning, theory classes, driving simulation training and practical driving instruction not in isolation, but rather as integrated elements of a pedagogically and didactically sophisticated system. Such integration enhances both the probability and the extent of synergy effects where different forms of teaching and learning are used. At the same time, depending on which forms of teaching and learning are combined with each other, integration promotes the transformation of declarative and implicit knowledge into diversely applicable action knowledge, and thereby enables the development of action routines to relieve the driver (Bredow & Sturzbecher, in press). A curriculum spanning both formal driving school training (i.e. theory classes and practical driving instruction) and further forms of teaching and learning used in novice driver preparation could facilitate the creation of reliable interfaces between the individual forms. Such comprehensive curricula are already in use in Finland and Iceland, for example.

The core prerequisite for a comprehensive curriculum is a set of training standards which establish a theoretical foundation not only for the system of training, but also for the system of testing. Such training standards must describe scientifically founded minimum requirements in respect of the expected extent of driving competence and thus define the desired learning results for driving school training. The test standards must be derived from the training standards and must reflect the latter's content. The development of training concepts relating to traffic perception and hazard avoidance, therefore, should also include the development of corresponding forms of testing (“traffic perception test”, see Chapter 2). The findings presented in this chapter with regard to the trainability of “traffic perception and hazard avoidance”, together with the described developments aimed at closer integration of such teaching/learning content into German driving school training, can be considered an important starting point: It must be possible to train those elements of content which need to be tested on account of their high safety relevance, and this training must actually be realised in a pedagogically and didactically appropriate manner in practice. Conversely, the development and trial implementation of traffic perception tests is able to contribute to the systematic training of competences relating to traffic perception and hazard avoidance in the driving school, as well as promoting the elaboration of suitable teaching and learning media.

[35] In a series of three experiments, Meir et al. (2010) were able to show that a combination of training elements in the sequence “Short active element” - “Instructional element” - “Longer active element” achieved the greatest learning successes.

*Jan Genschow, Bernd Weiße, Tino Friedel, Thomas Schubert, Conrad Teichert & Winfried Wagner and Sarah Malone & Roland Brünken*[36]

# Evaluation and further development of the theoretical driving test

## 1.12 Overview

Continuous evaluation of the theoretical driving test (TDT) is an important contribution to quality assurance and the definition of a scientific basis for the system of driving licence testing. The underlying evaluation concept, as well as the technical and organisational methods and procedures serving realisation, maintenance and further development of the TDT, are described in the "System Manual on Driver Licensing (Theory Test)" ("Handbuch zum Fahrerlaubnisprüfungssystem (Theorie)", TÜV | DEKRA arge tp 21, 2008) and have been implemented since 2008.

The experience gathered from practical use to date has led to further development of the system in terms of both working process organisation and the technical framework conditions. A proposal for restructuring of the responsibilities assigned to the involved bodies, for example, was elaborated by the Technical Examination Centres for Motor Vehicle Traffic (Technische Prüfstellen für den Kraftfahrzeugverkehr) and confirmed by the Federal/Regional Expert Committee "Driver Licensing and Driving Instructor Legislation" (Bund-Länder-Fachausschuss "Fahrerlaubnisrecht/Fahrlehrerrecht", BLFA-FE/FL) on 22nd March 2012. The associated organisational changes, which have ultimately led to faster processes and more efficient use of the available resources, were flanked by further development of the IT structures. For example, the administration and documentation program ("authoring system") which is used in the development and approval of new test items was adapted to the new organisational structure and expanded to allow also online access. Furthermore, changes in the form of test item visualisation (e.g. dynamic situation presentation) have provided starting points for technical further development of the system of testing. In this way, the "User Information System" ("AnwenderInformationsSystem – AIS"), which the publishers of teaching materials use for purposes of media design, has gained functions which enable test items and visualisations to be made available with minimum delay and without breaks in the media continuity. The functionality of the VICOM software used for test item visualisation, and likewise the scope of selectable content, has also been expanded substantially, taking into account the specific requirements of the training material publishers using the system. More detailed presentations of these changes can be found in the annual progress reports of the TÜV | DEKRA arge tp 21 working group for the years 2011, 2012, 2013 and 2014.

The following Chapter 4.2 discusses a number of analyses conducted during the report period to evaluate the parallel tests and individual test items which are currently in use by way of critical appraisal of the testing methods. Subsequently, Chapter 4.3 addresses the testing of new test approaches; it describes, for example, the didactic rationale behind the introduction of test items with alternative illustrations in October 2012 and presents corresponding research results. Chapter 4.4 is then devoted to the use of dynamic situation presentations in place of or alongside the usual static images of traffic situations. Dynamic presentations of developing situations promise a greater degree of congruence between

[36] Selected research results of the Chair of Empirical Educational Research at the University of Saarland are presented by Dr. Sarah Malone and Prof. Roland Brünken in Chapter 4.5.

the test demands and the actual demands to be mastered by drivers in real traffic (e.g. demands relating to information acquisition using the vehicle mirrors, visual scanning patterns, observance of relative speeds). This innovative instruction format was introduced (in combination with traditional multiple-choice questions) in 2014.

The aforementioned studies concentrate on developments in respect of the instruction format – the test items have to date been presented as multiple-choice questions. During the report period, however, ideas for innovative forms of response input were developed and tested at the Chair of Empirical Educational Research at the University of Saarland, first within the framework of a BASt project[37] and then under a commission from the TÜV | DEKRA arge tp 21 working group[38]. The project contents and selected results are presented in Chapter 4.5, together with conclusions relating to promising format developments. Chapter 4.6, finally, summarises some further topics for future research and development work in connection with the TDT.

## 1.13 Studies and measures within the framework of TDT evaluation

### 1.13.1 Continuous evaluation of test items and parallel tests

The evaluation concept for the TDT provides, among other things, for analysis of the various parallel tests with regard to their psychometric equivalence, and for analysis of the individual test items in accordance with certain test psychology quality criteria and content-related attributes. The evaluation results for each parallel test are summarised in a so-called "test sheet report", which contains not only information on the evaluation sample and its psychometric equivalence with other parallel tests, but also content descriptions and statistical values (e.g. question difficulty and discrimination) for the individual items belonging to the parallel test in the form of "test question profiles". On the basis of the evaluation results, test items with deficiencies are revised and/or optimised through the integration of innovative instruction formats (e.g. dynamic situation presentations).

By the end of the report period, a total of 166 parallel tests from the test sheet categories "Basic sheet – Initial licence", "Basic sheet – Extension" and "Supplementary sheet" for the driving licence classes "A", "A1", "B", "M", "C", "C1" "CE" and "D" had been evaluated. A total of 1,118 test items were considered in the course of these evaluations. It can be seen from Fig. 17 that more than two-thirds of the evaluated test items (67%, green bar; n=749) required no or only minor revision. For 29 per cent of the evaluated test items (yellow bar; n=324), optimisation was deemed expedient from the perspective of psychological testing, and the technical and content-related aspects of possible measures were discussed further and weighed up accordingly; realisation, however, is in many cases only feasible in the medium term (e.g. after development of innovative answer formats). The remaining 4 per cent of the evaluated test items (red bar; n=45) were found to display methodical deficiencies which – insofar as the test items concerned are not revised – call their further use in the TDT into question.

[37] The results of the BASt project "Test psychology and learning theory underpinning the test questions in theoretical driving tests, with special consideration of test question formats with image sequences (FE 82.326/2007)" were published as a BASt report (Malone, Biermann, Brünken & Buch, 2012).

[38] The corresponding research work is described in the project report "Validation of action-similar response formats for dynamic test items in driving licence tests" (Malone & Brünken, 2013a).

| | Test sheet category | Number of test items* | Items in need of optimisation (in %) | | |
|---|---|---|---|---|---|
| Basic sheet | Initial licence | 388 | 59 | 37 | 4 |
| | Extension | 265 | 55 | 41 | 4 |
| Supplementary sheet | A | 123 | 70 | 27 | 3 |
| | A1 | 92 | 71 | 26 | 3 |
| | B | 227 | 71 | 23 | 6 |
| | M | 85 | 72 | 25 | 3 |
| | C | 143 | 69 | 26 | 5 |
| | C1 | 73 | 69 | 23 | 8 |
| | CE | 102 | 69 | 30 | 1 |
| | D | 133 | 74 | 22 | 4 |
| **Overall pool of test items**** | | **1118** | **67** | **29** | **4** |

**Fig. 17: Test items evaluated to date and the determined need for optimisation**

* Number of different test items used for each driving licence class, i.e. without adjustment for multiple use for different classes.

** Total number of different test items evaluated, i.e. test items used several times in different parallel tests or in tests for multiple driving licence classes were only counted once.

The evaluation results presented here do not yet take into account the optimisation measures which have been elaborated in the meantime by the TÜV | DEKRA arge tp 21 working group and the working group "Theory test and test item development" (AG "Theoretische Prüfung und Aufgabenentwicklung"), because the optimised test items will only be subjected to renewed scientific evaluation after re-incorporation into regular test practice.

The optimisation recommendations and proposals derived from the initial scientific evaluation were taken up by the working group "Theory test and test item development" and – together with optimisation recommendations from other disciplines – formed the essential basis for continuous test item revision. This process of continuous revision was combined with necessary adaptation of the official catalogue of test questions in the light of changing demands (e.g. amendment of the Road Traffic Regulations). To date, over 84 per cent of the test items which – from the perspective of psychological testing – were found to be in need of immediate revision, have been revised accordingly, along with 42 per cent of the test items where optimisation was deemed expedient under the same premises. The test items concerned have already passed all the required stages of the approval process and have been added to the official catalogue of questions. Further revised test items are still at various stages of the approval process. The re-incorporation of revised test items into the official catalogue, and similarly the development of wholly new test items, are in this way contributing to continuous further development of the catalogue of questions.

### 1.13.2 Comparison of "conventional" and computer-generated image-based test items

With effect from July 2011, all conventional photographs and illustrations used in image-based test items contained in the official catalogue of questions were replaced with corresponding computer-generated images (see Fig. 18). In the past, considerable effort and expenditure had been required to produce or update the photographs used, as the relevant situations had to be found, or else simulated, in the real traffic environment. Computer-generated images, on the other hand, can be created much more efficiently; this simplifies both updating and later adaptation of the traffic situations to be depicted (Bönninger & Sturzbecher, 2005).

It can be recognised from Fig. 18 that the computer-generated images each present the same content as the conventional illustrations. The text elements of the test items (i.e. the question and the answer options) also remain identical in each case. Upon closer examination of the images, however, it becomes clear that there are certain fundamental differences between the original illustrations and their computer-generated reproductions (e.g. the representation of the dashboard in the computer-generated images, and differences in perceived quality between a “virtual” computer graphic and a “real” photograph). The extent to which these deviations may influence the processing of a given test item by the driving licence applicant cannot be clarified by way of a content-oriented analysis alone. Additional analysis of the candidates' response behaviour, i.e. the consideration of psychometric parameters and fit, thus seems appropriate as a statistical approach: If a computer-generated image is more or less suitable to depict a given traffic situation compared to a photograph, it can be presumed that a corresponding test item will be answered correctly by more or fewer candidates.

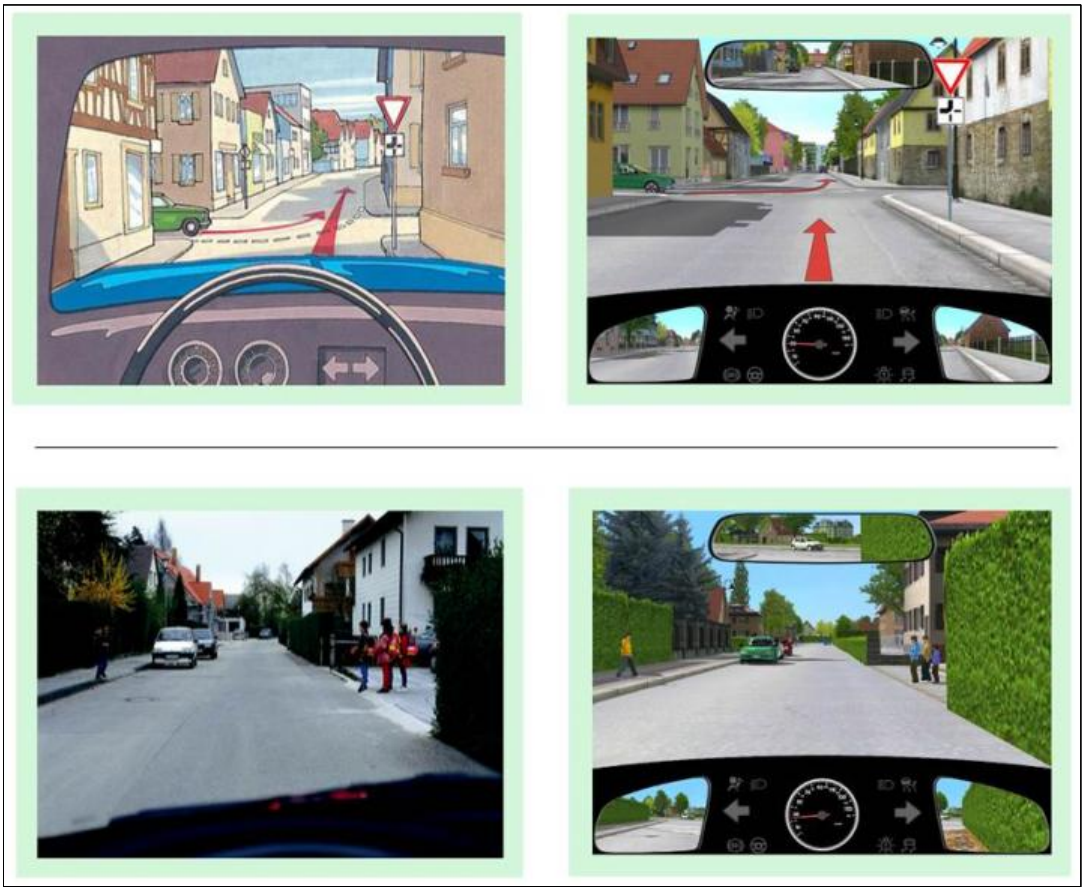


**Fig. 18: Original images (left) and computer-generated equivalents (right)**

Within the framework of an empirical study (Frisch, Burkert & Genschow, 2013), a total of 114 image-based test items were analysed by testing for statistically significant difference in the responses of driving licence applicants for selected pairs of test items (the same test item with conventional and computer-generated illustrations). The data for the comparison were collected from driving licence tests taken during three-month periods before and after the introduction of computer-generated images. The results are clear proof of the equivalence of the two types of images; there is evidently no disturbing influence to be attributed to the fundamental design differences (e.g. different perceived quality of computer-generated images and photographs) or new design elements (e.g. rear-view and side mirrors). In the few cases in which significant deviations in response behaviour were found, a content-oriented analysis of differences in the image – taking into account the offered solutions – supplied generally plausible explanations for the statistical differences. It was thus possible to derive scientifically founded proposals for optimisation of the test items concerned from the combination of statistical and content-oriented analysis, and furthermore to realise short-term implementation of these proposals.

### 1.13.3 Model calculations regarding optimisation of the system of assessment

Each test item in the official catalogue of questions carries a certain weighted score penalty (between 2 and 5 points) which is taken into account in the overall test result if the question is not answered correctly. The high degree of differentiation which is currently to be found in the weighting of test questions can be optimised from the perspective of psychological testing by selecting as many equally weighted questions as possible for the test construction (Lienert & Raatz, 1998). This would serve firstly to improve the practicability of test construction, because the differentiated weighting of individual test items currently influences the compilation of test sheets in compliance with the Examination Guidelines, in the sense that an identical maximum possible score must be guaranteed both in each subject area and for the test overall for all parallel tests used in the TDT. This means that, for formal reasons, it is not possible to actually use all the test items which are available in the official catalogue of questions. Furthermore, there seems to be little justification for the current weighting on the basis of differences in the traffic- or safety-relevant significance of the test content.

The necessity for modification of the system of assessment has already been presented and founded in more detail in various research reports on the TDT (Bönninger & Sturzbecher, 2005; Sturzbecher, Kasper, Bönninger & Rüdel, 2008; Genschow, Krampe & Weiße, 2011). Within the framework of a BLFA-FE/FL meeting on 19th September 2013, it was suggested that the expected influence of an optimised and simplified system of assessment should be studied and compared with the existing weighted system by way of model calculations, so as to provide an empirically founded basis for decisions on an alternative system of assessment. Starting out from the existing system of assessment[39] ("Status quo"), the following four alternative approaches were elaborated:

- Model 1: Full removal of the weighting; one point is awarded for each correct answer; the test decision depends on the total points score
- Model 2: Full removal of the weighting; one point is awarded for each correct answer; the test decision depends on the total points score, subject to the make-or-break criterion that two incorrect answers to questions from the subject area "Right of way" automatically lead to the candidate failing the test
- Model 3: Reduced weighting with one point awarded for each correct answer to general questions and two points awarded for each correct answer to questions from the subject area "Right of way"; the test decision depends on the total points score

[39] The individual test items carry a weighted score penalty of between two and five points for incorrect answers, and a class-specific maximum permissible penalty is defined to represent the points score which must be achieved to pass the test ("cut-off value"). Alongside the absolute number of penalty points, the regulation that the test is deemed to have been failed if two questions weighted with a score penalty of five points each are not answered correctly provides a further criterion for classification of the candidates as either "qualified to drive" or "not qualified to drive". As this error constellation always leads to failure of the test, it is also referred to as a "make-or-break criterion".

- Model 4: Full removal of the weighting; one point is awarded for each correct answer; the test decision depends on the total points score, subject to the make-or-break criterion that any incorrect answer to a question from the subject area "Right of way" automatically leads to the candidate failing the test.

On the basis of empirical test data (records of tests for various driving licence classes from the period from 11/2012 to 03/2013), it was determined how test outcomes could be expected to change, compared to the status quo, if assessment were to be based on each of the four alternative models (Cuvenhaus, Genschow & Sturzbecher, 2013). As a first step, the model calculations addressed possible changes in the pass rate for the TDT in connection with the four different assessment approaches (see Table 3). Depending on the specific number of items for different driving license classes, a "cut-off value" within the range from 90 to 92 per cent was applied. For the test for initial granting of a licence for vehicle class B, which comprises a total of 30 test questions, this value means that (under Model 1) at least 27 points (of the possible 30) must be achieved to pass the test. In the further cases, passing of the test was additionally subject to a make-or-break criterion (see Models 2 and 4), or else influenced by the higher weighting of certain questions with two points (Model 3).

**Tab. 3: Model calculations for TDT pass rates (driving licence class B)**

| System of assessment | | No. of tests passed (of a total of 94,518 tests) | Pass rate in per cent |
|---|---|---|---|
| **Model 1** | No weighting; 1 point per question | 68,320 | 72.3 |
| **Model 2** | No weighting; 1 point per question; two incorrect answers in the subject area "Right of way" as an additional make-or-break criterion | 68,188 | 72.1 |
| **Model 3** | Reduced weighting; 1 or 2 points per question | 66,465 | 70.3 |
| **Model 4** | No weighting; 1 point per question; any incorrect answer in the subject area "Right of way" as an additional make-or-break criterion | 64,095 | **67.8** |
| **Status quo** | Weighting; 2 to 5 points per question; two incorrect answers to "5-point questions" as an additional make-or-break criterion | 65,052 | 68.8 |

As can be seen in Table 3, there is only a slight change in the average TDT pass rate for driving licence class B when a new system of assessment is applied. The smallest change results in case of full removal of the current weighting in combination with introduction of a make-or-break criterion according to which the candidate fails if any question relating to right of way is not answered correctly (Model 4; cell highlighted grey in the column "Pass rate"). The pass rate is here reduced by only one per cent compared to that for the present weighted test assessment with two to five points per question ("Status quo"). The remaining assessment approaches (Models 1, 2 and 3), by contrast, lead to an increase in the pass rate, i.e. more candidates than in the past would pass the TDT; the greatest increase by 3.5 per cent could be expected when applying Model 1, i.e. unweighted assessment without a make-or-break criterion.

As a further means to predict the changes which could be expected from an optimised system of assessment, it was determined how the classification of candidates as either "qualified to drive" or "not qualified to drive" would change. In Table 4, it is shown for each of the four assessment models, how many candidates would receive a different overall test result (i.e. "test passed" or "test failed") if a new system of assessment were to be applied, i.e. the shifts in classification between "not qualified" ("fail") and "qualified" ("pass") compared to the status quo.

**Tab. 4: Model calculations for shifts in the classification as qualified/not qualified to drive in the TDT (driving licence class B)**

| System of assessment | | ←→ Total | | Pass → Fail | | Fail → Pass | |
|---|---|---|---|---|---|---|---|
| | | N | % | N | % | N | % |
| **Model 1** | No weighting; 1 point per question | 3,828 | 4.05 | 280 | 0.30 | 3,548 | 3.75 |
| **Model 2** | No weighting; 1 point per question; two incorrect answers in the subject area "Right of way" as an additional make-or-break criterion | 3,696 | 3.91 | 280 | 0.30 | 3,416 | 3.61 |
| **Model 3** | Reduced weighting; 1 or 2 points per question | 2,777 | 2.94 | 682 | 0.72 | 2,095 | 2.22 |
| **Model 4** | No weighting; 1 point per question; any incorrect answer in the subject area "Right of way" as an additional make-or-break criterion | **5,099** | **5.39** | 3,028 | 3.20 | **2,071** | **2.19** |

Note: The percentage changes each refer to the base value of 94,518 tests; N designates the absolute numbers of cases in which the classification changed; →, ← and ←→ indicate the direction of change between "Pass" and "Fail" or vice versa, or else the overall changes in either direction.

It can be seen in Table 4, that the proportion of changed classifications is smallest for Model 3. Of the total of 94,518 tests analysed, only 2.94 per cent (N=2,777) resulted in a different classification compared to the present system of assessment. This can be attributed above all to the fact that a large number of candidates who failed in the past would be classified as "qualified to drive" under Model 3 (approx. 75.4 per cent of the changes in classification are from "fail" to "pass"). The situation is similar for Model 1 (3,828 changes in classification, of which an even higher proportion of 92.7 per cent refer to changes from "fail" to "pass") and for Model 2 (3,696 changes in classification, of which 92.4 per cent refer to changes from "fail" to "pass"). The largest proportional change in classification (cells highlighted grey in the column "←→ Total") results from implementation of Model 4, where 5.39 per cent of the candidates (N=5,099) are classified differently compared to the current system. Contrary to Models 1, 2 and 3, however, the reclassification is here predominantly from "pass" to "fail" (i.e. 59.4 per cent of the reclassified candidates; cells highlighted grey in the column "Fail → Pass"). This is evidently due to the fact that the make-or-break criterion is stricter under Model 4 (the majority of the candidates who passed under the present system, but would no longer pass according to this new model, had answered a question from the subject area "Right of way" incorrectly).

If the findings relating to the pass rates and the changes in classification between "qualified to drive" and "not qualified to drive" are viewed together, it is initially noticeable that the smallest change in pass rates compared to today's status quo is to be expected with Model 4: The test pass rate for driving licence class B is here only 1 per cent lower, whereas it deviates much more significantly (by up to 3.5 per cent) with the other models. The changes with regard to the classification of candidates as either "qualified to drive" or "not qualified to drive", on the other hand, were greatest with Model 4. According to the criteria of this model, 3.2 per cent of the candidates who currently pass the TDT (driving licence class B) are classified as "not qualified to drive" – without there being any appreciable change in the overall pass rate. The future application of this model would evidently affect above all those candidates who have to date benefited from the fact that the test can still be passed when a question from the subject area "Right of way" is answered incorrectly. In other words, the make-or-break criterion would be applied more strictly – and from the perspective of test didactics more consistently: Only those candidates who have fully and reliably understood the rules of right of way are able to pass the TDT. Under the remaining models (1, 2 and 3), on the other hand, there are benefits for those candidates whose corresponding knowledge is incomplete. As shown by the above analyses, this can also be expected to correlate with a corresponding increase in the pass rate. On the basis of the model calculations performed, it is thus Model 4 – unweighted scoring for all test questions, in combination with a make-or-break criterion – which is recommended for optimisation of the current system of assessment. Given its pedagogical-didactic foundation, application of this model would lead firstly to more a precise discrimination of qualified and non-qualified candidates – with practically no impact on the present pass rate. At the same time, the practicability of test construction would be improved significantly.

### 1.13.4Comparison of training and test contents

The TDT is a criterion-referenced test which serves to assess whether a driving licence applicant, within the framework of his driver training, has attained learning objectives which are deemed decisive for the granting of a driving licence. As the contents of driving school training and driving licence testing are anchored in different statutory regulations, namely the Learner Driver Training Ordinance (Fahrschüler-Ausbildungsordnung, FahrschAusbO) and the Driving Licence Regulations (Fahrerlaubnis-Verordnung, FeV), it is possible that individual aspects of the training have no immediate equivalent in the test contents, and vice versa – despite the essential correspondence between training and test contents. Already in 2009, a method for comparative analysis of the training and test contents was developed as a means to identify those topics which are not reflected congruently in both driver training and subsequent testing, and applied in the first instance to the subject area “Right of way and traffic rules”. For a valid analytical comparison, however, it is not sufficient to consider merely the overarching subject areas defined in the FahrschAusbO and FeV, respectively; the examination must include also comparative appraisals of the specific contents of individual test items in each subject area, on the one hand, and the practice and monitoring exercises in the corresponding training materials used in driving school training (in particular the text books offered by commercial publishers), on the other hand.

During the period of the present report, the comparative analysis of training and test contents was continued for the subject area “Environment protection” and the procedure for processing and presentation of the results was expanded. Besides content relating to driving licence class B, for example, the evaluation now includes also content relevant for the driving licence classes A, C and D. In addition, a closely detailed examination of specific content elements determined how often individual aspects were covered in the analysed media. This permits more precise statements on the content status of an analysed subject area, for example the number of different aspects of content which are currently reflected in the test questions for a particular subject area (“differentiation”), the relative extents to which certain aspects of content presented in the analysed training materials are currently addressed in actual test questions (“weighting”), or the specific aspects of training media content which are either represented or not represented in the available test questions (“congruence”). The results are intended to help the branch experts involved in the optimisation of novice driver preparation to better judge the sufficiency of contents relating to a particular subject area, and to set priorities for the further revision and expansion of the available pool of test questions (“subject-area-specific revision”). The comparative analyses are to be continued for further subject areas.

## 1.14 Development of variants of computer-generated static illustrations

### 1.14.1Starting points from the perspective of test didactics

The introduction of computer-generated illustrations in 2011 was an important prerequisite for the realisation of further methodical and didactic benefits of a computer-based

TDT. The development process has since been continued with the elaboration of so-called “illustration variants”, which were developed with the objective of forcing candidates to actively analyse the depicted content where traffic situations are presented in image form in traditional multiple-choice questions. The measures which had already been taken up to this point to prevent manipulation and to disturb the schematic, merely superficial memorisation of answer patterns consisted in randomisation of the order in which test questions and answer options are presented. With regard to the image-based questions used in the TDT, this serves to hinder the rote learning of scenario/solution pairs (e.g. “That is the image where I need to tick the first and third answers”).

The desirable necessity to analyse the actual content of a depicted situation, however, cannot be attained through random variation of the test arrangement alone. It is to be expected that, as before, the images will simply be linked superficially to the correspondingly required behaviour in the form of correct answer options (e.g. “When that image is displayed, I must wait”, “When that image is displayed, I have right of way”, and so on), as studies on memory formation assign a high level of memorability to images. Standing (1973), for example, showed a group of test persons 2,000 slides on each of five consecutive days. Even two days later, they were still able to recognise 73 per cent of this total of 10,000 images (Engelkamp, 2004). Alongside the outstanding ability to recall images, it is similarly easier to memorise images compared to terminology, i.e. a person can apparently remember the image of an object better than the name or designation of that object (Engelkamp, 2004); this is known as the “picture superiority effect”.

How – against the background of these findings from basic research – can the driving licence applicant be brought to analyse specifically the content of a traffic situation depicted in test questions? Deeper involvement with the purpose of a question (rather than merely schematic recognition of the required answer) can probably only be achieved if the image presented in the test situation is not one which is already visually known to the candidate, but instead shows a traffic situation which, although comparable in terms of content to the illustrations encountered during test preparation, bears no visual similarity. This assumption is supported by the aforementioned studies on the recognisability of images, according to which the degree of similarity between a previously seen image and an unknown image is decisive for the success of recognition (Engelkamp, 2004). It is to be expected, therefore, that the presentation of visually dissimilar images during test preparation, on the one hand, and in the actual driving licence test, on the other hand, would lead to the candidate paying less attention to the pictorial details of the image and consequently to concentration on the pertinent traffic-relevant information (e.g. “Which traffic rules apply in this situation?”).

The similarity between known and unknown images may refer to either the visual or the conceptual information contained in the image. The visual information in an image concerns the form and colour of the depicted objects (design of individual elements of the image), whereas the conceptual information represents the content or context of the illustration. Engelkamp (2004) points out that it is not only visual similarity, but also visual complexity which influences the recognition of images which are otherwise identical in respect of their conceptual content. Thus, if several images which are similarly complex and furthermore conceptually comparable (as is the case in the TDT) are presented, it can be assumed that modification of the visual information will reduce recognisability. With regard to the initially defined objective, therefore, it appears promising to hamper visual recognition through variation of the form and colours of elements depicted in the traffic environment, and in this way also to promote active consideration of the conceptual image content.

### 1.14.2 Empirical study on the recognisability of traffic-related illustrations

To examine how image variation may affect recognisability in the context of the TDT, the TÜV | DEKRA arge tp 21 working group conducted a corresponding empirical study using computer-based images (Friedel & Rüdel, 2011). The first intermediate results of this study were presented in the innovation report for the years 2009 and 2010; in the meantime, the study has been concluded. The goal of the study was to determine whether variation resulted in images no longer being recognised when viewed again, and, if so, from which degree of variation this effect could be observed. This was intended to supply empirically founded knowledge on the image elements which must be varied and the extent of variation necessary to reduce recognisability without altering the perceived content of an image (Friedel et al., 2011). In the following, the methodical approach and the essential results of the study are described. More detailed information can be found in the corresponding research report “Recognition of varied static depictions of traffic situations” (Friedel & Rüdel, 2011).

In preparation for the empirical study, a set of 30 computer-generated static images was created, each of which was very similar in content to the illustrations of current test questions. A representative cross-section of different traffic situations was selected (e.g. both urban and rural environments). Subsequently, the 30 initial images were modified in respect of certain design elements, such as the types and numbers of buildings, the road surface, the type and scope of vegetation, or the types and numbers of road users. The individual image elements were varied systematically and to differing extents, so as to be able to determine the influence of variation (independent variable) on the recognisability. In addition to the recognisability of the original images (first dependent variable), which was intended to be as low as possible, the study also sought to determine the extent to which illustration variants were perceived as conceptually or contextually similar to the corresponding original image (second dependent variable); ideally, this perceived similarity was to be as high as possible. Evidence of such correlations would confirm the expediency of variation in the sense of the aforementioned didactic objective.

A total of 30 persons took part in the study (20 women and 10 men with an average age of 31 years). Seventeen study participants already held a driving licence (on average for the past eight years) and had gathered an average of approx. 7,000 kilometres of driving experience. The study required each participant to attend two appointments, namely a “Learning” session (first study appointment) and a “Recognition” session (second study appointment). During the first appointment, the study participants were asked to view 30 computer-generated static images of different traffic situations (“original” illustrations). To promote active consideration of the conceptual image contents, the participants were required to indicate the behaviour which would be appropriate in the given traffic situation and to give corresponding explanations. On the next day (second appointment), the study participants were again shown 30 images; this time, 15 of the images were unchanged (i.e. they corresponded to the “original” images from the previous day), whereas the other 15 had been varied systematically. For each of the 30 images, the participants specified whether they had recognised the image (measurement of the first dependent variable “Recognisability”). In addition, they were asked to assign their situation assessments from the previous day to a current image (measurement of the second dependent variable “Content similarity”).

The results of the study are illustrated in Fig. 19, with the first dependent variable “Recognisability” plotted on the Y axis and the second dependent variable “Content similarity” on the X axis. It can thus be seen how – in the sense of the didactic objective – the different degrees of variation in the image influenced recognisability and the perceived content similarity. The ideal effect of a variation would be a maximally high value for the variable "Not recognised" and at the same time a maximally high value for the variable "Perceived as similar"; this is represented by the green area of the diagram. Each of the symbols used in the diagram stands for a certain degree of variation compared to the original image.

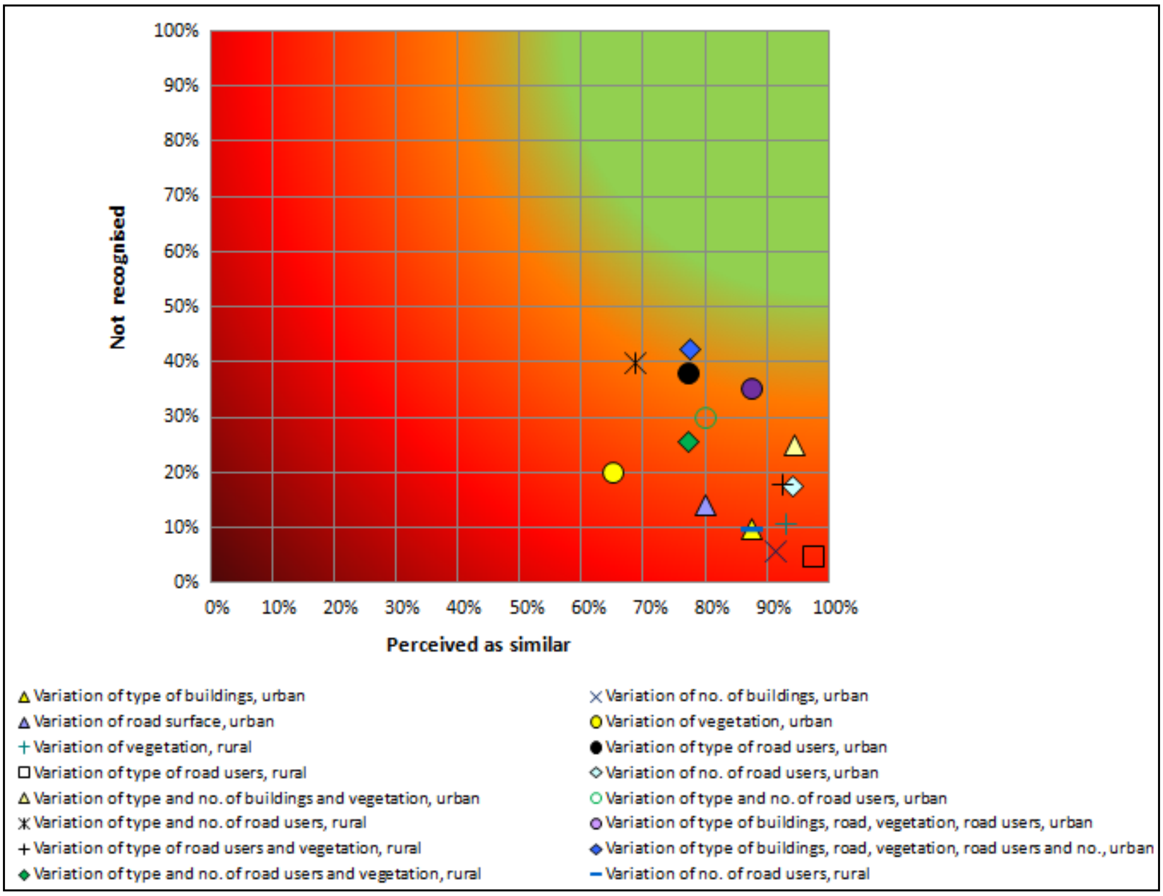



**Fig. 19: Impact of variation on recognisability and content similarity**

The diagram shows clearly that the combined variation of several image elements, in particular, produces the desired effect (blue diamond). Variation of the type of road user also leads to distinct reduction of the recognisability (black dot). These statements apply at least with regard to depictions of an urban traffic environment, whereas it is the combination of varied types of road user and varied numbers of road users which achieves a relatively low degree of recognisability in the case of rural traffic situations. Furthermore, the results indicate that practically all illustration variants possess a high level of content similarity.

The results of the study show that variation of the images used in the TDT hinders their recognition, and that the illustration variants are perceived as equivalent in terms of conceptual content, irrespective of their deviation from the original illustration.

### 1.14.3 Process rules for the elaboration of illustration variants and their integration into the TDT

Taking the aforementioned empirical findings, it is possible to derive process rules for the creation of illustration variants for use in the TDT. These rules serve to guarantee that the illustration variants available for a particular test question can be considered equivalent in terms of content. Each illustration variant is created on the basis of an “original” image and the corresponding task as defined in the official catalogue of test questions. To ensure content equivalence between the original illustration and the (usually five) variants, the conceptual elements of the image (i.e. the positioning of the traffic situation on the basic screen, the positions of the viewer's own vehicle and all other vehicles or persons directly involved in the traffic situation, task-relevant traffic signs, signals and road markings) are preserved in each of the variants. Fig. 20 presents an example of an original image (here the question with the official reference number 1.1.02-040-M) and the five corresponding illustration variants (V1 to V5); the red highlighting is intended to indicate that the illustrations are interchangeable.

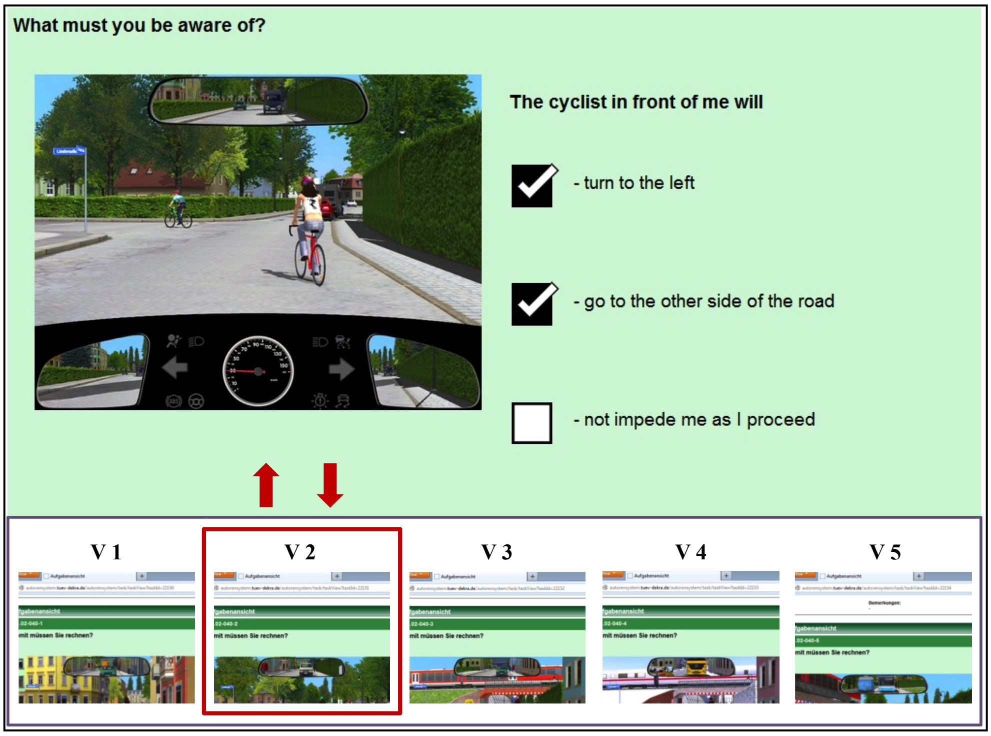



**Fig. 20: Original test question with instructions and answer options (above), together with five corresponding illustration variants (below)**

For variation of the depicted road users, clearly distinguishable colours are used; reference is also made to these colours in the text instructions and in the answer options. The vehicle types may be varied, insofar as this does not alter the conceptual basis of the task for the test candidate. Other road users who are not connected with the task addressed by the test question can always be depicted with a different appearance, placed in a different part of the image or even omitted completely. Variation of the surrounding buildings, vegetation and any infrastructure which is irrelevant to the question (e.g. the road surface, insofar as this is of no significance for the test content) is realised in various ways. Generally speaking, the aim is always to “redevelop” the surroundings. Such variations in the street architecture serve to represent typical differences in urban landscape. Finally, the weather and time of the day (visibility conditions) are also varied, provided this does not modify the test question in respect of content.

The process for the creation of illustration variants is flanked by a comprehensive system of quality assurance. Illustration variants are designed by the TÜV | DEKRA arge tp 21 working group and checked by the working group "Theory test and test item development" before release. As soon as test questions with illustration variants are used in the TDT, their equivalence is verified empirically, taking into account the relevant psychometric parameters, within the framework of continuous evaluation (see Chapter 4.3.4).

To guarantee equal test conditions for all candidates, only the "original" images are published in the official transport ministry gazette ("Verkehrsblatt") to announce new image-based questions with illustration variants. The original images, however, are not used to illustrate questions in the actual test; only the unpublished illustration variants are used for this purpose. The illustration variants are thus unknown to the candidates until the time of their test. The original images, on the hand, are to be found in relevant teaching materials and can thus be used for test preparation. In all cases, it is only the illustration which is replaced; the formulated questions and the answer options remain unchanged.

The unpublished illustration variants are stored on the candidate computers. When a test sheet is assigned, corresponding illustration variants are selected at random and presented to the candidate with the test questions. Information on the presented variants is saved in the candidate data record for evaluation purposes.

### 1.14.4Evaluation of variants of computer-generated static illustrations

As test items with illustration variants have been used in regular driving licence testing since 1st October 2012, it is now both possible and necessary to realise a systematic evaluation of such test items. One of the essential questions in connection with the newly implemented instruction format obviously concerns the equivalence of the illustration variants: When creating illustration variants, the challenge is to vary those object attributes which are irrelevant for processing of the test item, without altering the actual content of the question (see Chapter 4.3.1). To ensure that this is achieved in the desired manner, it must be tested empirically, whether all variants of a test item are equally suitable to present the traffic-specific correlations and thus to assess attainment of the learning objective. If, indeed, only object attributes which are of no significance for the processing of the actual test demands are varied, then the pattern of candidate responses to a given question should remain practically unchanged, irrespective of the illustration variant presented.

Following introduction of the first illustration variants in regular testing, empirical data from the test records of subsequent driving licence applicants were analysed to determine whether the different illustration variants derived from an original image ("variant quintet") were actually to be deemed equivalent and thus equally suitable as test item instructions. To this end, a selection of 13 different test items, i.e. a total of 65 illustration variants, was examined with regard to statistically and practically significant deviations in the responses of test candidates within each variant quintet (Frisch, Teichert & Genschow, 2013). The results showed that no practically significant differences were to be found in respect of the test item difficulty – all instances of a variant quintet were thus solved or not solved equally frequently. The same is for the most part also reflected in the "attractiveness" of the different answer options, i.e. the frequency with which the correct or incorrect answers were selected within a variant quintet. In only one of the 65 cases analysed was there any statistically and practically significant deviation compared to the responses of those candidates who had been presented one of the other four variants. Through comparisons of the test content, however, it could be shown that this was attributable to task-specific design attributes, and the illustration variant concerned was immediately revised accordingly.

On the basis of the available study results, it can be concluded that the test content addressed by a particular test item can generally be transplanted quite successfully into different traffic environments, without this affecting the actual intention of the test item. The primary objectives of the illustration variants, namely to counteract superficial rote learning, to encourage more intense preparation and to promote a situation-independent understanding of traffic-related contexts, thus appear credible from the perspective of test methodology. It is furthermore evident that the analyses represent a valuable source of hints regarding optimisation of the test items. Such pointers are deemed particularly important, because the introduction of illustration variants means not only that the diversity of the situations depicted in image-based questions increases, but also that additional design options and equivalence demands must be taken into account in test development. It was thus recommended in the aforementioned study report, that the study approach outlined here should in future be applied to all test items with illustration variants within the framework of continuous evaluation.

## 1.15 Development of instruction formats with dynamic situation presentations

### 1.15.1 Framework conditions for the introduction of dynamic instruction formats in the TDT

Both the evidently more realistic representation of actual traffic demands and the self-explanatory nature of dynamic situation presentations are to be mentioned as general benefits of this innovative instruction format in the context of the TDT. The enhanced realism stems from the fact that, through the use of a computer as the test medium, the inherent dynamics of road traffic can now be addressed as a new aspect of test design. Even though the illustration is necessarily simplified, it becomes possible, for example, to depict relative speeds, successively developing elements of a situation, spatial relationships (taking into account also the use of side and rear-view mirrors) or even hazard cues which are only visible temporarily. The presentations are at the same time self-explanatory in that traffic circumstances can be visualised and, where necessary, modulated in their complexity without requiring long textual instructions to explain the situation. The reduction of verbal situation descriptions serves to minimise content-independent influences on a candidate's mastering of the test demands (e.g. reading competence, language understanding).

To enable the introduction of dynamic situation presentations into the regular TDT, the current choice of instruction formats (comprising textual and illustrated questions) has been expanded to include also short video sequences with a duration of around 15 seconds. The traditional multiple-choice style has been retained; even where the situation is presented as a dynamic scenario, the candidate is still offered the usual answer options in the sense of attractors and distractors. Regarding the use of test items with dynamic situation presentations, the Federal/Regional Expert Committee “Driver Licensing and Driving Instructor Legislation” (BLFA-FE/FL) adopted the following framework conditions at its meeting on 28th/29th September 2011:

- For each test sheet, two of the current test items are to be replaced with new items using dynamic scenarios.
- No restrictions apply regarding the subject areas to which test items using dynamic scenarios may belong.

- The position of the test items within the test sheet is to be determined at random, as for the current items, and each new test item is to be marked as a “dynamic test item” by way of a suitable icon.
- The video sequences can be viewed up to five times before switching to the screen in which the actual question is presented. The candidate is to be informed as to how many times the sequence can be viewed (again).
- A proposal with regard to the score weighting of each test item is to be made, as in the past, by the experts from the working group “Theory test and test item development”.

To guarantee an adequate number of parallel tests, a total of 52 such test items were developed and officially approved before introduction of the new instruction formats. The TÜV | DEKRA arge tp 21 working group established the technical prerequisites for integration of the new test items into the existing test environment by way of a comprehensive update of the test system software. Fig. 21 below shows an example of the screen views presented to the candidate in the test application.

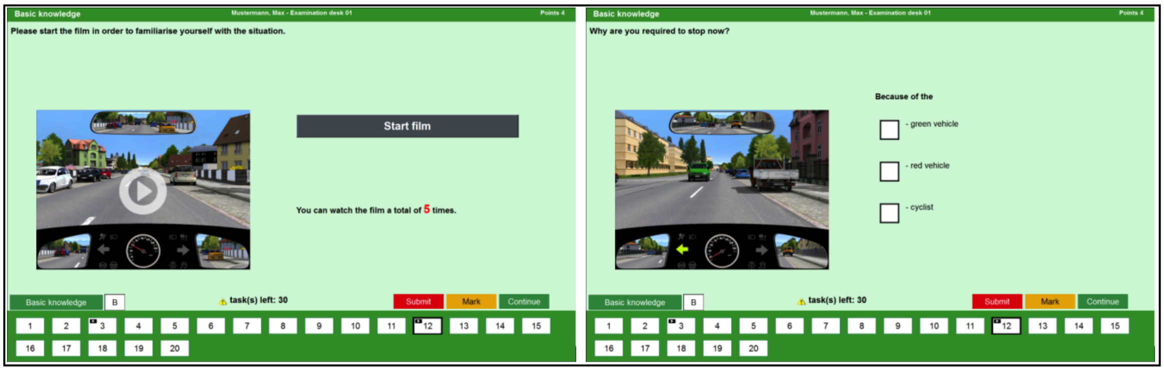



**Fig. 21: Example of a test item with dynamic situation presentation**

As can be seen on the left-hand screen in Fig. 21, the candidate is initially informed that it is possible to view the video sequence up to five times. This information is updated after each viewing to indicate the remaining number of views, and the candidate is furthermore able to move on to the actual question. The corresponding screen view is to be seen on the right in Fig. 21; it comprises the last image of the dynamic scenario, together with the test question and the answer options. In the regular TDT, the test items with dynamic situation presentations are placed at random between positions 1 and 20 of the basic test sheet. From the candidate's point of view, the test items concerned are marked with a corresponding symbol (white triangle on a black background). The proven layout of the test application thus remains essentially unchanged, even after the introduction of dynamic scenarios.

### 1.15.2 Examples for the development and practical testing of new test items

The content for test items incorporating dynamic scenarios was either developed from scratch or else based on the content of existing test items which had been classified as in need of revision within the framework of continuous evaluation. Where new content was to be developed, the relevant literature was analysed in combination with accident statistics, in order to identify novice-specific driving deficits and novice-typical accidents which could be taken as a starting point for a corresponding test item. The results of these analyses, however, were for the most part still not detailed enough: Descriptions of typical novice deficits, for example, are often too vague (e.g. “insufficient horizontal scanning of the traffic situation”) and the accident classifications in accident statistics too general (e.g. “Accident in longitudinal traffic in which the driver strayed off the road without a conflict with other road users”) to serve as templates for test demands and depictable traffic situations. It is rather the case that appropriately founded details must be added to a potentially suitable situation. Suitable sources are here specific accident data and experiential reports (e.g. from driving instructors). In the following, the development process is illustrated with two examples of new test items. First of all, the procedure for completely new test content is explained (test item example “Accidents in longitudinal traffic”). The second example then refers to the modification of an existing test item which was found to be in need of revision (test item example “Tailgating”).

*Test item example “Accidents in longitudinal traffic”*
A test item relating to a hazardous situation in longitudinal traffic was developed on the basis of a known accident black spot in Saxony (see Fig. 22).

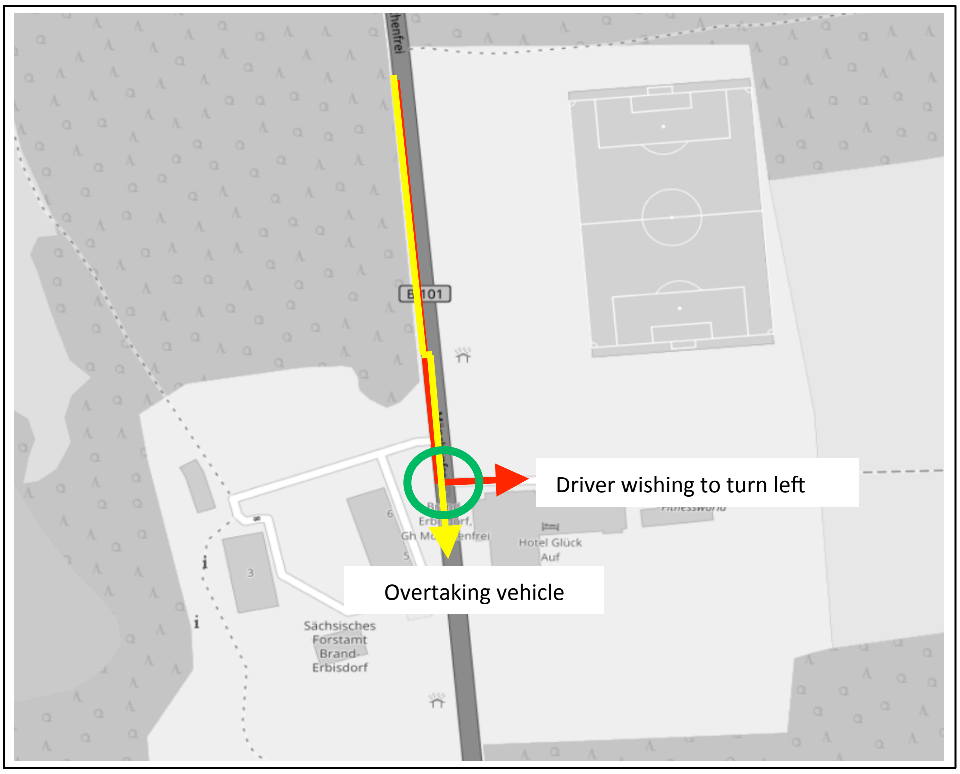



**Fig. 22: (Former) accident black spot in Saxony**

Fig. 22 above depicts the potentially hazardous driving lines of two road users: The primary-class road here runs through a section of forest and a small hamlet, to which a hotel also belongs. This straight section of road is the only opportunity to overtake in the wider vicinity. At the time the test item was elaborated, the permissible maximum speed when passing the hamlet was 100 km/h, and there was no restriction with regard to overtaking. Hazardous situations arose when drivers wishing to turn left towards the hotel (red arrow) reduced their speed as they approached the hamlet. This could lead to accidents if, at the same time, another driver (who was perhaps not familiar with the local surroundings) emerged from the forest just ahead of the previously concealed turn-off point and saw a brief opportunity to overtake the apparently slower vehicle (yellow arrow) which was actually preparing to turn left.

A 15-second video was created with the VICOM editing software to depict the described risk situation, and a test item was formulated with corresponding answer options (see Fig. 23). In the elaborated dynamic scenario, which is presented from the driver's perspective (“ego-vehicle”), the candidate sees that his vehicle is approaching a small hamlet. There are no traffic signs which indicate a speed limit, overtaking restrictions or the start of a built-up area, as is also the case at the real accident black spot. As the situation unfolds, the ego-vehicle slows down and the left turn indicators begin to flash, evidently because the driver wishes to turn into a driveway. Due to the deceleration of the ego-vehicle, a further vehicle slowly closes up from behind (white vehicle in the mirrors), and it can be seen that this vehicle is being overtaken by a third vehicle (red vehicle in the mirrors). In this situation, the candidate should recognise that he should allow the red vehicle to complete its overtaking manoeuvre – irrespective of the fact that such behaviour is in contravention of § 5 (4) sentence 4 of the Road Traffic Regulations – so as to avoid causing an accident. To this end, he must recognise that the red vehicle in the mirrors is overtaking and then judge its speed. From the estimation of the distance and speed of the overtaking vehicle, it must be derived that safe turning is only possible after the red vehicle has passed.


**Fig. 23: Elaborated test item prototype with dynamic situation presentation (final image) relating to accidents in longitudinal traffic**

*Test item example "Tailgating"*

On the basis of an existing test item with exclusively textual instructions (see Fig. 24), a new item was elaborated to present the same traffic situation in the form of a dynamic scenario. This test item had previously been earmarked for revision in the course of continuous evaluation, which had identified a need for corresponding optimisation: The existing test item placed high demands on the reading competence of the candidate; furthermore, certain formulations in the answer options led to the correct and incorrect answers appearing especially plausible or unlikely (e.g. "… as soon as this is possible").

**What do you do if a driver is following you closely on the autobahn and urges you constantly by flashing his headlights to clear the overtaking lane?**

- [x] You move over to the right lane as soon as possible
- [ ] By braking slightly you urge the following car to keep a greater distance
- [ ] You immediately cut in between cars travelling at a safe distance in the right-hand lane

**Fig. 24: Test item (withdrawn) with exclusively text-based instructions**

Using the VICOM editing software, the original instructions were visualised by way of a traffic situation in which the ego-vehicle is overtaking a number of other vehicles on the motorway and, to do so, is currently driving in the leftmost lane. During the presented scenario, a fast-moving vehicle approaches from behind in the same lane. This vehicle closes up to less than the safe following distance, flashes its headlights, drops back a little and then closes up once more; Fig. 25 shows the final image of prototype developed for this test item. A comparison of the old and new instruction texts shows clearly how the question is now much shorter and the unwanted solution hint has been eliminated.


**Fig. 25: Elaborated test item prototype with dynamic situation presentation (final image) relating to tailgating on the motorway**

The test demands are changed as a result of the dynamic visualisation, in the sense that the candidate must now himself recognise that he would violate the safe following distance between two vehicles if he were to move over immediately after passing the van (incorrect answer B), and that it is thus necessary to overtake also the blue car (correct answer A). Furthermore, it must be realised that the already insufficient following distance of the tailgating vehicle would be reduced further by any braking on his own part (incorrect answer C). This test item is a good example for the significantly improved self-explanatory nature of dynamic situation presentations.

*Practical testing of new test items*

The “System Manual on Driver Licensing (Theory Test)“ (TÜV | DEKRA arge tp 21, 2008) specifies that test item prototypes which deviate from the conventional formats (text only, image/text combination) must be evaluated in practical trials. Regarding the new test items with dynamic situation presentation, corresponding studies were conducted from October 2009 to March 2012 (Friedel & Rüdel, 2010): Within the framework of the regular TDT, driving licence applicants were asked to answer a number of prototype questions based on the new instruction format – in each case after completing their actual test, but before receiving their result. The study sample embraced a total of around 20,000 participants; the age and gender distributions within the sample corresponded to those of TDT candidates overall. Approximately two-thirds of the participants passed the test; this also corresponds to the typical pass rate. The study sample was thus not subject to any systematic sampling error and permits the study results presented below to be generalised to refer to all driving licence applicants.

Fig. 26 shows the range and distribution of difficulty indices (p) for 30 prototype test items with dynamic situation presentation. The test item prototypes presented to candidate for a Class B driving licence (red columns) were spread over a difficulty range from 0.04 to 0.93. Candidates who were applying for an initial licence were less able to solve the test items compared to applicants for an extended licence. This corresponds to expectations, as the dynamic instruction format mirrors real traffic demands to a high degree and thus favours candidates with traffic experience. The apparently small proportion of test items with a high frequency of correct solution can be explained by the fact that the items presented were prototypes and thus entirely unfamiliar to the candidates at the time of the study. It is likely, therefore, that the difficulty indices will fall once the test items concerned are addressed specifically within the framework of test preparation (e.g. in driving school training, in teaching/learning media).

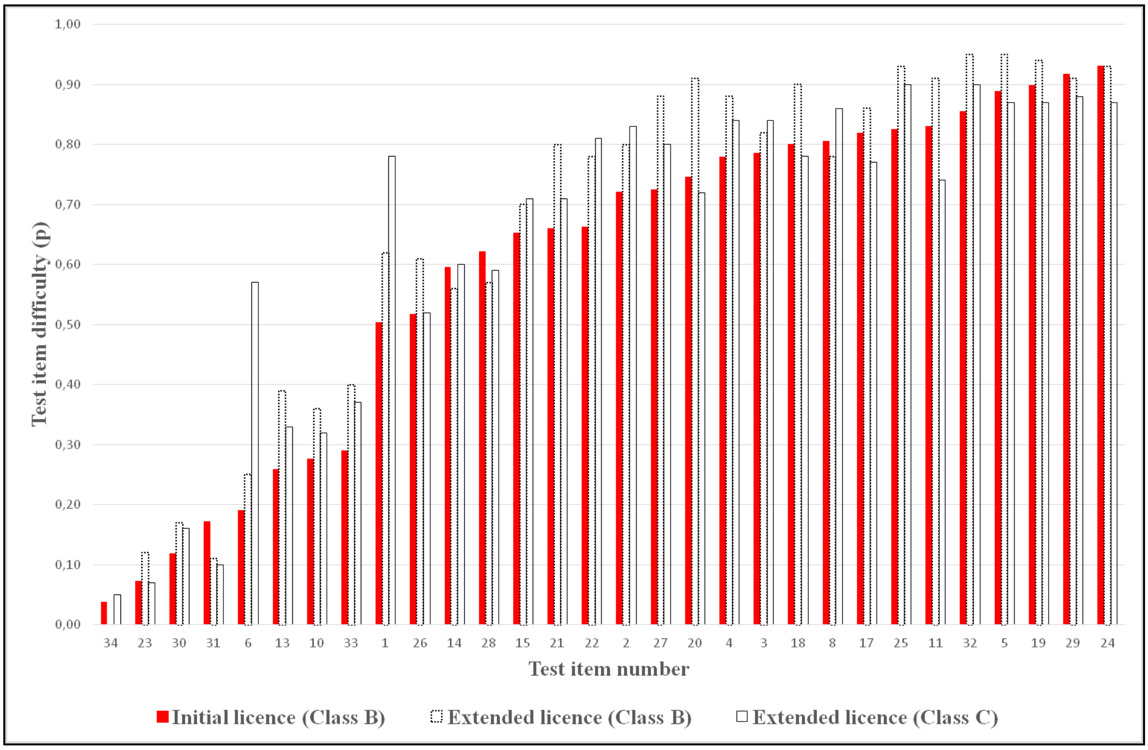



**Fig. 26: Difficulty indices (p) of test item prototypes according to test type (initial driving licence, extended driving licence)**

The study participants were also asked to provide various socio-demographic data, for example their level of school education or computer experience, so as to permit an analysis of possible performance differences between sub-groups (e.g. between candidates with different school education backgrounds or with greater or lesser computer experience). In the main, no correlations – or at most very minor correlations – were revealed. The innovative instruction format is thus evidently not a source of special difficulties for certain candidate groups; it is rather that equal test conditions apply for all candidates. The process of practical testing outlined here can be deemed proven as a means of empirical test item evaluation and should also be used in the future in the development and testing of new test item formats.

### 1.15.3 Overview of the test items with dynamic situation presentations elaborated to date

In the following, the 52 test item prototypes which had been approved up to the end of the report period are described in more detail, firstly with regard to their classification according to the subject areas of the official catalogue of test questions, and secondly in respect of the visualised test content.

*Classification of test items according to subject area*

The 52 test items with dynamic situation presentations are split equally between the two subject areas “Hazard perception” and “Behaviour in road traffic” (see Table 5). It would naturally also be possible to make meaningful use of dynamic situation presentations to illustrate test items relating to other subject areas (e.g. “Traffic signs”, “Right of way”). However, as the added value compared to conventional instruction formats is most evident in the subject areas of “Hazard perception” and “Behaviour in road traffic”, these areas have been treated first.

**Tab. 5: Classification of test items according to subject area**

| **Subject areas in accordance with the official catalogue of test questions and the corresponding numbers of test items** | | | |
|---|---|---|---|
| **Hazard perception** | 26 | Motorways | 1 |
| | | Special traffic situations | 12 |
| | | Road and weather conditions | 1 |
| | | Speed | 2 |
| | | Overtaking | 6 |
| | | Behaviour towards pedestrians | 4 |
| **Behaviour in road traffic** | 26 | Turning and reversing | 5 |
| | | Use of road lanes | 7 |
| | | Special traffic situations | 3 |
| | | Vehicles with flashing blue and yellow lights | 1 |
| | | Public transport and school buses | 1 |
| | | Overtaking | 4 |
| | | Passing stationary vehicles and obstacles | 3 |
| | | Traffic lights and illuminated signs | 2 |

*Classification of test items according to the visualised traffic environment*
Dynamic situation presentations can be constructed and described systematically on the basis of certain categories derived either from typical driving demands or from the applicable legal framework (e.g. road traffic regulations). One common categorisation is the distinction between traffic situations “within built-up areas” and “outside built-up areas”. Traffic situations within built-up areas, for example, are usually characterised by a relatively high traffic density and complexity, and by the presence of vulnerable road users; such situations can be differentiated further by considering the character of the locality (urban, rural). Situations outside built-up areas, on the other hand, are characterised by the high speeds of the individual participants and thus demand appropriate judgements with regard to relative speeds, acceleration and braking. The type of road (motorway, country road) is often considered for further differentiation in the case of traffic situations outside built-up areas. Table 6 classifies the 52 test items with dynamic situation presentations according to the visualised traffic environments.

**Tab. 6: Classification of test items according to traffic environment**

| Visualised traffic environments and the corresponding numbers of test items | | | |
|---|---|---|---|
| Traffic situations within built-up areas | 24 | Urban locality | 22 |
| | | Rural locality | 2 |
| Traffic situations outside built-up areas | 28 | Country roads | 13 |
| | | Motorways | 15 |

The reference to traffic environments is one possible form of categorisation and description for the content of dynamic situation presentations, and can serve as a basis for refinement and expansion by way of further demand-relevant categories in the future. A first instance of such systematisation is described in the following section.

*Classification of test items according to driving actions and situational reasons for action*
Each of the 52 test items presents a certain driving action, or else a correspondingly planned action (e.g. flashing turn indicators on the dashboard), viewed from the perspective of the driver. These driving actions (see Table 7) can usually be associated with certain situational circumstances, which must thus be interpreted correctly by the test candidate (“occasions for action”). Through understanding of the (planned) driving action and identification of the relevant situational reasons for action, the candidate is able to decide whether an action can be performed without violating traffic regulations and/or without endangering his own or other vehicles.

**Tab. 7: Driving actions and reasons for action**

| **(Planned) driving actions of the ego-vehicle** | **Situational reasons for action** | | | | |
|---|---|---|---|---|---|
| | Obstacles, traffic rules, traffic signs | (Concealed) hazards | Other vehicles changing lanes | Other vehicles overtaking | |
| Overtaking/passing | 4 | 6 | 3 | 9 | **22** |
| Turn off to the left | - | 4 | - | 3 | **7** |
| Turn off to the right | - | 4 | - | - | **4** |
| Stay in driving lane | 7 | 5 | 1 | 6 | **19** |
| | **11** | **19** | **4** | **18** | **52** |

Table 7 shows the different driving actions and the associated situational reasons for action which have been visualised in the dynamic situation presentations elaborated to date. It can also be seen that hazard awareness is a frequent aspect of the depicted driving actions. The hazards concerned are either visible only momentarily during the course of the scenario and then concealed once more (e.g. children running onto the road behind a parked vehicle), or else they are not immediately apparent (tram which approaches from behind while the driver is waiting to turn left and is thus only visible in the mirrors). The planned actions “Overtaking and passing” and “Stay in driving lane” are similarly frequent topics. The test items with dynamic situation presentations thus already cover a broad spectrum of traffic situations in different environments, and place a particular focus on potential accident situations. As they currently still use the traditional multiple-choice design, however, they remain subject to certain methodical limitations (e.g. the often necessary explicit mentioning of hazard cues in the answer options, or the possibility of merely guessing a correct answer); this deficit can only be overcome through the development of innovative answer formats.

## 1.16 Development and testing of innovative answer formats

### 1.16.1 Starting points

Since nationwide introduction of a computer-based TDT, it has become possible to exploit the benefits of computer-based methods for competence assessment within the framework of driving licence testing, for example through the use of more situationally realistic presentation formats[40] (e.g. dynamic scenarios) and more directly action-referenced answer formats (e.g. recording of the reaction time). The BASt research report “New exercise formats in theoretical driving tests” (Malone, Biermann, Brünken & Buch, 2012) presented innovative test item formats developed specifically for this purpose, together with results from empirical studies to assess their suitability from the perspectives of psychological testing and teaching/learning theory. Particular consideration was given to the use of those dynamic presentation formats which, given the close correspondence between the imposed test demands and the associated perception processes in the natural driving context, seemed to be the most promising. The test item formats elaborated by the project (knowledge questions, judgement tasks, measurements of reaction time) were each implemented in two presentation formats (static, dynamic) and analysed by way of expert-novice comparisons to determine their criterion validity (i.e. their effectiveness as a means to distinguish between “experts” who have already gathered driving experience and “novices” who to date possess relatively little driving experience).

The results show that action-referenced test items which serve to measure reaction times (similar to the “Hazard Perception Test” used in Great Britain) display a particularly high criterion validity compared to knowledge questions and judgement tasks. A longitudinal study to assess the performance development of learner drivers during driving school training revealed furthermore that, over time, the learner drivers improved in their handling of all three test item formats. Even so, it was only in respect of items which queried declarative knowledge content that the learners achieved the same performance level as experienced drivers already during their training. In the case of items geared to reaction

[40] The term “presentation format” used in this chapter refers exclusively to aspects of the visualisation of traffic situations (e.g. static, dynamic) in test items. By contrast, the term “instruction format” which is used elsewhere in the present report is broader in scope and encompasses all aspects of the test item instructions (e.g. the textually formulated question, instructions regarding answer input; see also TÜV | DEKRA arge tp 21, 2011).

time, which serve rather to assess action-specific procedural knowledge, the novices were able to improve their performance over time, but were nevertheless unable to match the results of the experts. Against the background of the study results, the use of dynamic situation presentations in the TDT was recommended in principle: The items with dynamic and static illustrations displayed comparable criterion validity, but dynamic situation presentations are to be considered ecologically more valid.

*Objectives and theoretical background of the project "Validation of action-referenced answer formats for dynamic test item formats in driving licence testing"*

The results of the aforementioned BASt project and further related studies (Malone & Brünken, 2013b) suggested that test item formats with a more concrete reference to actual driving actions could achieve greater criterion validity. Consequently, a follow-up project was commissioned by the TÜV | DEKRA arge tp 21 working group to examine the possible benefits of test items which implemented not only an innovative presentation format (dynamic instead of static situation presentations), but above all also innovative answer formats with a concrete action reference. As described in more detail in the next sections, two studies were conducted within the framework of the project "Validation of action-specific answer formats for dynamic test item formats in driving licence testing" to assess the validity of answer formats with different levels of action reference by way of cross-sectional (expert-novice comparison) and longitudinal analysis. It was to be determined whether the potential to distinguish between experts and novices increased in line with the action reference of the answer format. The study design was to permit identification of the answer formats which displayed the highest criterion validity (i.e. the clearest difference between experts and novices) and those with the highest instructional validity (i.e. the best modelling of the novice learning curve).

Further foundation for the thoughts on the possible importance of action reference for criterion validity can be derived from empirical findings on novice-specific driving competence deficits and from the results of expertise research. The accident risk for a young novice driver, for example, is appreciably higher compared to older and more experienced drivers. The sharp decrease in novice risk within the first few months of solo driving, however, is indication that expertise acquisition in the domain of driving progresses relatively quickly in this period. Through studies with expert-novice comparisons, it is possible to determine abilities which are already well-developed in experienced drivers, but still only rudimentary in inexperienced drivers. Differences revealed between the two groups allow conclusions to be drawn as to the deficits which could be a cause of accidents. Numerous studies have tested individual methods to measure such differences in levels of expertise (e.g. Huestegge, Skottke, Anders, Muesseler & Debus, 2010; McKenna & Crick, 1994a), but only very few have compared different methods (Malone, 2012; Scialfa, Borkenhagen, Lyon & Deschènes, 2013). Such methodical comparison should thus be taken into account when investigating differently action-referenced answer formats.

As a prerequisite for comparative analyses of the significance of different levels of action reference, the task instructions and contents ("stimuli") should remain identical in each case as far as possible, and only the answer format should be varied. When designing the stimulus material, advantage should be taken of the benefits which empirical educational research assigns to the new media in respect of competence acquisition and assessment (Hartig & Klieme, 2007). Alongside economic considerations, these benefits include also the possibility to enhance objectivity, reliability and validity as the three primary quality criteria of psychometric test methods (cf. Jurecka & Hartig, 2007). As far as the TDT is concerned, the possibility to integrate dynamic presentations of traffic scenarios is especially promising as a means to improve quality. Research on learning with new media has been able to show that the use of dynamic materials has beneficial effect where certain conditions apply: According to Höffler and Leutner (2007), dynamic presentation formats are superior to static illustration where the animated objects are representative rather than decorative elements, the more realistic the animation, and where procedural rather than declarative learning content is to be acquired. Malone and Brünken (2013b) demonstrated that these findings can also be applied, in part, to achievement testing. Novice drivers benefit from dynamic situation presentations to a greater extent than experienced drivers. This seems to indicate that dynamic presentations aid the learner driver's comprehension of the situation and are thus able to reduce measuring errors, which in turn raises the reliability and validity of the test items concerned.

On the other hand, the studies conducted by Malone and Brünken (ibid.) also show that, as long as the text instructions and answer options are left unchanged, there is little to be gained from merely replacing the static illustrations currently used in the TDT with dynamic presentations of the traffic situation concerned. Although the authors provided evidence that learner drivers benefit from dynamic presentation formats, the multiple-choice questions used were found to be unsuitable to distinguish between learners and experienced drivers, whether with dynamic or with static presentation of the traffic situation. Since the material for the study was based on existing test items with static illustrations, with dynamic situation presentations being created for incorporation into test items which otherwise remained unchanged in terms of language and content, it can be assumed that the material contained only situations which could be adequately illustrated with a static photograph or drawing. As the complex demands of real driving could not be reflected meaningfully on the basis of the existing task contents, it seemed more expedient to examine the superiority of experienced drivers in conjunction with situation presentations and answer formats which conform better to the cognitive demands placed on the driver of a motor vehicle. This assumption is supported theoretically by findings from expertise research, which show that experts outperform novices all the more clearly where the set task is characteristic for the domain in question (e.g. Glaser & Chi, 1988 and, especially for sports domains, e.g. Hodges, Huys & Starkes, 2007; Thomas, Gallagher & Lowry, 2003). Subsequently, therefore, the studies in connection with dynamic presentation formats were broadened to consider also the use of innovative answer formats with concrete reference to the action demands of real driving.

### 1.16.2 Cross-sectional and longitudinal empirical studies

It is desirable for a test to be able to measure learning and experience gains as precisely as possible. In the context of the current project, the means chosen to analyse learning progress over the course of theoretical and practical driving school training, and then through initial solo driving experience, was a longitudinal study (see *Study 2* below) wherein the performance of the participants (“novices”) was recorded at several points during their driving school training. In advance of and in preparation for this longitudinal analysis, a preliminary study was conducted to select test item and answer formats which permit the differentiation of experts and novices on the basis of performance (see *Study 1 – “Preliminary study”* below). A further point to be addressed within the framework of the study was the extent to which the steps already implemented by the TÜV | DEKRA arge tp 21 working group to optimise test items using the traditional multiple-choice format (e.g. revision of the text instructions, answer options and visualisations) had contributed to improved test item quality. To this end, the original test items were compared with the revised versions.

*Study 1 ("Preliminary study"): Testing of action-referenced formats by way of cross-sectional studies*

Three test item formats with demands which differed in respect of their specificity (low, medium, high) for the domain "driving" were chosen for an expert-novice comparison. It was anticipated that the experts would achieve better results than the novices across all test item formats. At the same time, the experts were expected to outperform the novices all the more clearly with increasing specificity of the answer format used:

- The closed, multiple-choice format (in the following also referred to with "MC") was assumed to possess a low action reference. The task for the test person – after watching a short video – was initially to decide whether or not the driver must reduce his speed on account of the traffic circumstances. If the test person decided that speed reduction was appropriate, he was subsequently asked to specify his reasons for this decision. Brief descriptions of four possible reasons, one of which was the correct solution, were offered as answers to this additional question.
- An answer format with measurement of the reaction time had already been tested in earlier studies and was here used once more with the same animated traffic scenarios. With this format, which was assigned medium specificity, the task for the test person was to react by pressing a key on the keyboard (space bar) as soon as a reason for the driver to reduce his speed was recognised in the computer-animated traffic scenario (Reaction "Hand"). It was recorded firstly whether the task was solved correctly, i.e. whether the person reacted within a predefined time window (hits) or else did not react if no information cue appeared (correct rejections). In addition, the reaction time was measured as a further dependent variable.
- To further enhance the specificity of the answer format, a third test variant replaced the reaction by way of the space bar on the keyboard with activation of a foot pedal (Reaction "Foot"). In similar fashion to braking when sitting in a vehicle, the task for the test person was to react to appropriate stimuli by pressing the pedal with his right foot.

Fig. 27 provides an overview of the test design, the tasks to be solved by the test persons and the variables measured.

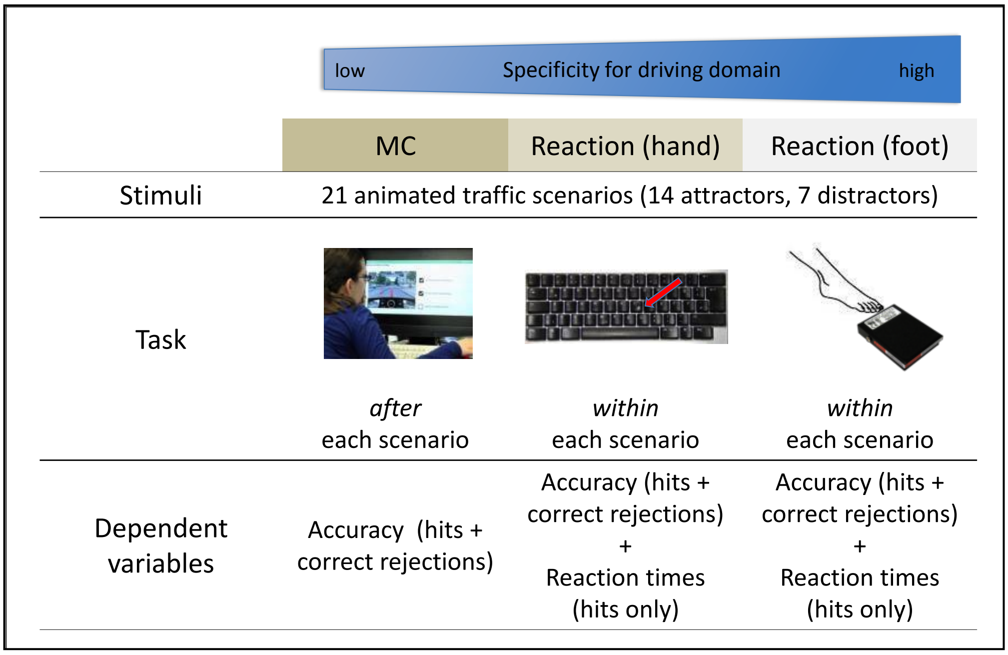



**Fig. 27: Test design in Study 1 ("Preliminary study")**

The specificity of the test items used in the study was varied by way of the action reference of the answer format. The three test item formats thus differed only in respect of the answer format used, whereas the traffic scenarios presented remained the same. A total of 21 dynamic situation presentations which were deemed suitable for use in conjunction with the three different answer formats were selected from previous studies. The research plan necessitated a 2x3 design with the factors “Expertise” (novice driver or driver with several years of driving experience) and “Answer format specificity” (low, medium or high). Each test person handled the set tasks by way of either the MC answer format, the reaction format using a keyboard or the reaction format using a foot pedal. A total of 116 persons participated in the study, of whom 60 were novices (learner drivers). The persons recruited for the expert group were exclusively drivers who had possessed a Class B driving licence for more than two years and had already gathered at least 5,000 kilometres of solo driving experience.

To test the hypotheses, a two-way analysis of variance was performed with the two-level factor “Expertise” and the three-level factor “Answer format specificity”. The dependent variable was the average number of correctly solved test items. The analysis of variance revealed a significant main effect for the factor “Expertise” ($F(1,109) = 10.25$, $p = .002$, $\eta_p^2 = .09$); overall, the experts solved more items correctly than the novices. There was also a significant main effect for the factor “Answer format specificity” ($F(2,109) = 3.92$, $p = .02$, $\eta_p^2 = .07$). The lower the specificity of the answer format, the more test items were solved correctly. No significant interaction was found between the two factors “Expertise” and “Answer format specificity” ($F(2,109) = 1.20$, $p = .31$, $\eta_p^2 = .02$). The descriptive data nevertheless show that the experts outperformed the novices to a noticeably lesser extent in the case of the reaction answer format with the foot pedal compared to the other two answer formats.

Discriminant function analyses were performed for the individual test item formats to classify the study participants as either experts or novices on the basis of their success in solving the set tasks. Where the format provided for the recording of a reaction time, it was also tested whether the reaction time – as a second expertise-dependent measure of performance alongside the number of correctly solved test items – aided the correct classification of experts and novices. Overall, the format “Reaction (Hand)” produced the best predictions (73% correct classifications). The MC format, with its single predictor “Number of correctly solved test items”, was only marginally less effective (71% correct classifications). The format “Reaction (Foot)”, however, was found to be unsuitable (54% correct classifications). It was furthermore revealed that the dependent variable “Reaction time” had no positive influence on classification by way of the answer format “Reaction (Hand)”. In fact, consideration of the sole predictor “Number of correctly solved test items” actually improved the rate of correct classification compared to the previous analyses with both predictors: 90 per cent of the novices were identified.

*Study 2: Testing of action-referenced formats by way of longitudinal studies*

The second study aimed to determine the extent to which the elaborated test items, and thus answer formats with different levels of action reference, are able to mirror the learning progress of a novice driver. It was assumed that the experts would achieve better average results than the novices, but also that the novices would display clearer improvements over time compared to the experts. Given the results of the preliminary study (“Study 1”) with regard to the classification of experts and novices, it was decided that the reaction format using the foot pedal was to be disregarded. Furthermore, it was to be tested whether the already implemented optimisation of the original multiple-choice items of the TDT had achieved any increase in validity. For these test items, the following three levels of revision were distinguished for the purposes of the study:

- Five exclusively text-based items and ten test items with computer-generated static illustrations were selected from the official catalogue of test questions and used in unchanged form (“Old test items”).
- For a further 15 test items (all of which made use of illustrations), the original language of the instructions and/or answer options had been revised in the course of optimisation (“Classic revision”).
- Another 15 test items were optimised by incorporating dynamic situation presentations. These test items addressed the same content as the illustrated test items of the aforementioned revision level, but their final static image was in itself not sufficient to answer the question correctly, as important information cues which were displayed in the course of the dynamic scenario were no longer visible in the final image (“Innovative revision”).

Alongside these three sets of 15 test items, the 21 items in multiple-choice format from Study 1 (“MC preliminary study”) and the 21 items which required a keyboard response from Study 2 (“Reaction item”) were also included in the study. Overall, therefore, the longitudinal study considered five different types of test item. While the test items of the first three sets corresponded exactly in terms of content, those of the remaining two sets were taken from merely similar areas of content.

For Study 2, a 2x3x5 design with repeated measurements of one factor was chosen. The two factors “Expertise” (expert or novice) and “Answer format” (old test items, classic revision, innovative revision, MC preliminary study or reaction item) were not measured repeatedly. The repeatedly measured factor “Time” was assigned three levels. Consequently, the test material was presented to each study participant on three occasions (measuring points MP1 to MP3) at intervals of between three and four months. The participants were allocated one of the five types of test item at random at the time of the first measurement, and then retained this type of test item for the two subsequent measurements. While the novices were at the beginning of their driving school training at the time of the first measurement (MP1), the second measuring point (MP2) coincided approximately with the TDT, and the third measurement (MP3) was positioned shortly before or shortly after the granting of a driving licence. A total of 236 persons was counted at the beginning of the longitudinal study, of whom 147 were novices. The results obtained at the time of the first measurement are presented in graphic form in Fig. 28.

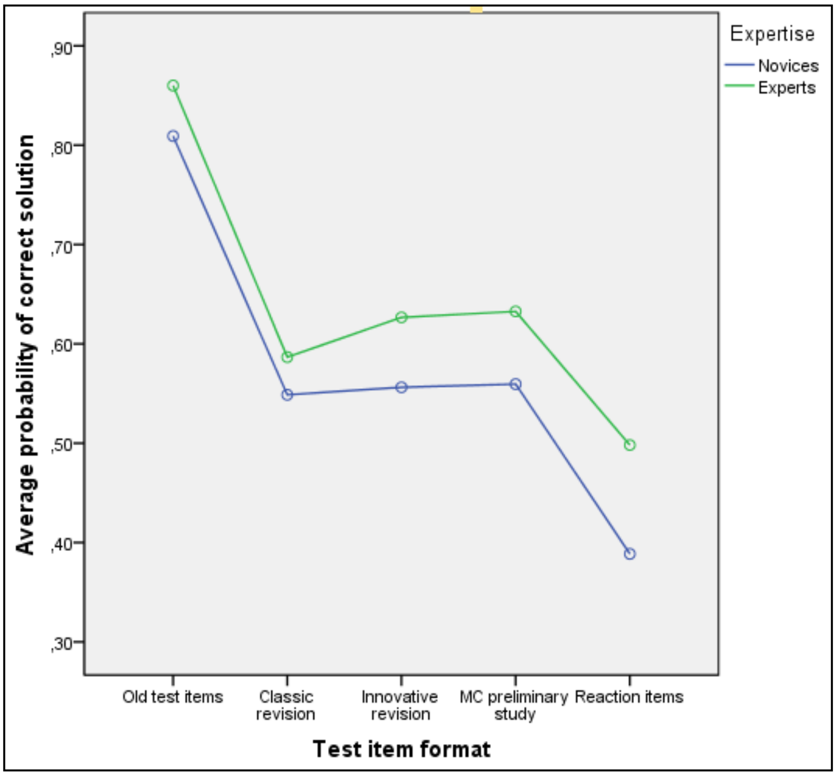



**Fig. 28: Results of the first measurements (measuring point MP1)**

For inferential statistical analysis of the data at MP1, a two-way analysis of variance was performed with the two-level factor “Expertise” and the five-level factor “Test item format”. The dependent variable was the average probability of solving the test items correctly. The analysis of variance revealed a significant main effect for the factor “Expertise” ($F(1,224) = 11.18$, $p = .001$, $\eta_p^2 = .05$); overall, the experts solved more items correctly than the novices. There was also a significant main effect for the factor “Test item format” ($F(4,224) = 39.34$, $p \leq .001$, $\eta_p^2 = .41$). The average probability of solving the test items correctly thus varied appreciably between the different test item formats. “Old test items”, for example, led to significantly better results than test items in any of the other formats. Furthermore, the test items of the format “Reaction item” were significantly more difficult that the other test items. No significant interaction could be found between the two factors “Expertise” and “Test item format” ($F(4,224) = .33$, $p = .86$). The descriptive data nevertheless show that the distinction between experts and novices was not equally successful across all test item formats. While only small differences were revealed between experts and novices for the test item formats “Old test items” and “Classic revision”, the gap between experts and novices was much clearer with the three remaining test item formats.

For the analysis relating to performance development over time, it was only possible to use the data of those persons who had participated at all three measurement sessions. Table 8 shows the distribution of these participants between the different combinations of factors. For inferential statistical analysis of the data, a three-way analysis of variance was performed with repeated measurements for the three-level factor “Time”. The further factors used in the analysis were the unrepeated factors “Expertise” (two levels) and “Test item format” (five levels).

**Tab. 8: Samples of the longitudinal study**

| N = 113 | | Answer format | | | | |
|---|---|---|---|---|---|---|
| | | Old test items | Classic revision | Innovative revision | MC preliminary study | Reaction items |
| Expertise | Novices (learner drivers) | n = 12 | n = 8 | n = 16 | n = 12 | n = 10 |
| | Experts (Driving licence >2 years) | n = 16 | n = 10 | n = 14 | n = 8 | n = 7 |

The analysis of variance with repeated measurements revealed no main effect for the factor “Expertise” ($F(1,103) = 2.40$, $p = .13$). Neither of the two expertise groups achieved significantly better results than the other group over time and across all test item formats. A statistically significant main effect was found, however, for the factor “Test item format” ($F(4,103) = 34.97$, $p < .001$, $\eta_p^2 = .58$). The “Old test items” were much simpler and the “Reaction items” were much more difficult that the other types of test items. There was furthermore a statistically significant three-way interaction between the factors “Time”, “Expertise” and “Test item format” ($F(8,194) = 2.16$, $p = .04$, $\eta_p^2 = .08$), as illustrated in graphic form in Fig. 29.

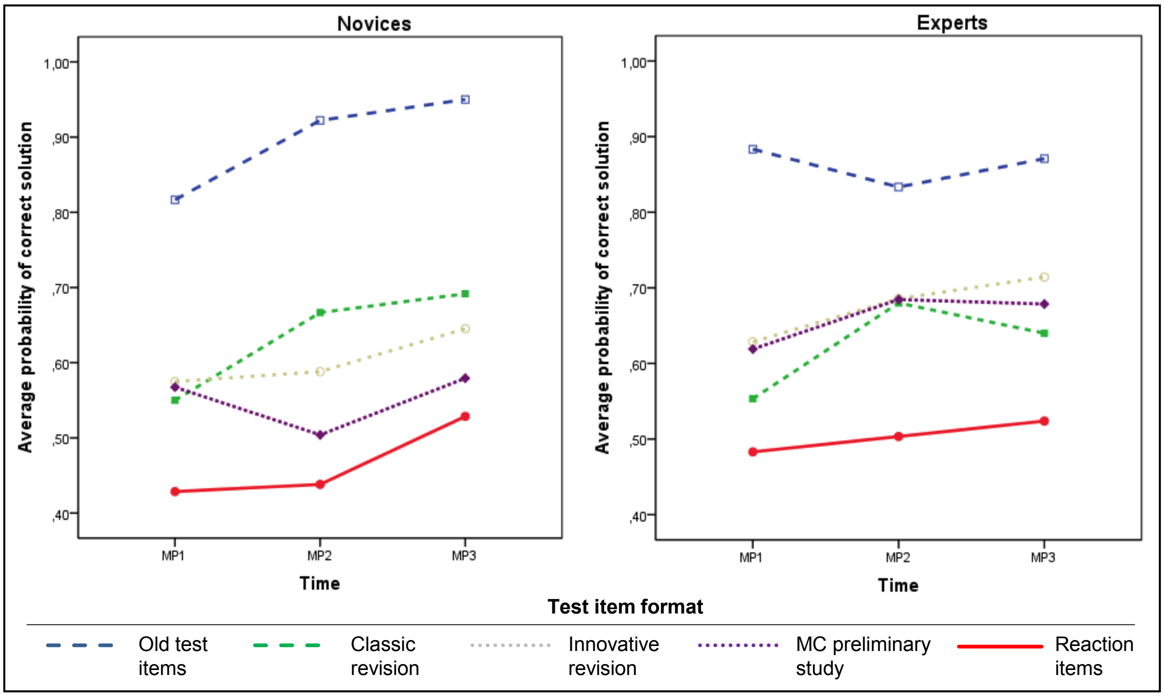



**Fig. 29: Illustration of the interaction Time*Expertise*Test item format**

The performances of the experts and novices developed differently over time for the different test item formats. The novices improved across all test item formats, with the exception of the format “MC preliminary study”, where there was initially a drop in the results between MP1 and MP2. It is furthermore conspicuous that, for the test items with static presentation formats (“Old test items” and “Classic revision”), the novices improved to a greater extent between MP1 and MP2 than between MP2 and MP3. Where the test items were designed with dynamic presentation formats (“Innovative revision”, “MC preliminary study” and “Reaction items”), on the other hand, the greater improvement in performance was between MP2 and MP3. The experts, by contrast, recorded little improvement over time. Minor linear increases were determined only for the test item formats “Innovative revision” and “Reaction item”.

### 1.16.3 Discussion of the results

The objective of the research project presented here was to identify the test item formats and specific design attributes which are most suitable for use in the TDT. It was shown that the “old test items” of the TDT, which used either text only or else text illustrated with a static computer-generated image, posed a very low level of difficulty and were furthermore less able to distinguish between experts and novices. On the other hand, the novices were able to improve their results over the course of driving school training, which can be taken as evidence for the instructional validity of the test items. The initial optimisation measures implemented by the TÜV | DEKRA arge tp 21 working group, which have concerned the formulations of the task instructions and answer options, have proved successful in the sense that they have raised the demand level of the test items – without detracting from the instructional validity.

Clear differentiation between experts and novices, and thus proof for the criterion validity of the test items, is only revealed after optimisation in the form of dynamic situation presentations. In such test items, a previously static image was not simply replaced with an animation. The transience of the animation format was also utilised: Relevant aspects of the traffic situation only appear temporarily, and are then no longer visible in the final image of the scenario. It must be noted, however, that these test items were less suitable to reflect the knowledge gains of novice drivers during the phase of declarative knowledge acquisition. The performances of the novices only improved appreciably once they progressed to practical driving. Consequently, there is here a certain incongruity between the instructional and criterion validity, which may be an indication that the learning objectives of driving school training lack adequate criterion validity. This assumption is supported by the results for the reaction-based test items and corresponding items in multiple-choice format from Study 2. Both formats distinguished between experts and novices, but failed to mirror the knowledge gains from theoretical driving school training.

A further important result regarding reaction time measurements and their suitability for future use in action-specific test items in the TDT is the fact that no increase in criterion validity was attained through the use of a foot pedal as the input device. The use of foot pedals for reaction-based test items cannot be recommended at present. Furthermore, it is to be noted that, for the reliable classification of novices without driving experience, it was sufficient to determine whether or not the person concerned had reacted to an information cue within a predefined critical time window. The speed of the reaction within this critical window was irrelevant for the quality of classification.

On the basis of the two cross-sectional and longitudinal studies, the reaction-based test items requiring a keyboard input generally appear to be the most promising of the innovative answer formats tested for further development of the TDT. The longitudinal studies showed that candidate performance improves as learning progresses, i.e. once the acquired contents of the theoretical training are supplemented with practical driving experience. If such test items are used in actual driving licence testing, it can thus be expected that driving licence applicants would choose to take their test at a later date than is the case at present, as this would increase the probability of success. Last but not least, test item formats with dynamic situation presentations could be a meaningful complement to the practical driving test, as they permit behaviour assessments in situations which are difficult to realise in real traffic (e.g. particular hazards, different weather conditions). Further studies will be required, however, to confirm the results obtained to date and to refine the knowledge gained. If such test item formats are to be integrated into a future driving licence test, a larger pool of test items must first be created along the lines of the test items used in this research project. Subsequently, trials with larger samples would also be necessary to evaluate their specific characteristics.

## 1.17 Outlook

Another important step in the methodical further development of the theoretical driving test in Germany was taken during the last year of the period covered by the present report: Since 1st April 2014, dynamic situation presentations have for the first time been used as a further, innovative instruction format for the classic multiple-choice-style test items. The implementation and regular use of test items with dynamic situation presentations in the TDT must now be followed up with assessment of the methodical quality of these test items within the proven process of continuous evaluation.

A number of important questions are already emerging and must be answered by the pending evaluation. Given the varying complexity of the depicted traffic situations (number of road users to be taken into account, weather conditions, etc.) and of the demands to be met in a particular case (action decision, detection of a hazard cue), for example, it would be interesting to determine any systematic correlations to relevant statistical values (difficulty, discrimination). Questions targeting possible correlations between the successful completion of test items with dynamic situation presentations and the use of computer-based teaching/learning media or the number of practical driving lessons taken to date appear similarly pertinent. Last but not least, the possible benefits of essentially self-explanatory animated scenarios, as opposed to text-based instructions to describe a given situation, could contribute to test success being determined above all by traffic-related competences, to the exclusion of content-alien factors such as the reading competence of driving licence applicants. To be able to answer the arising questions, it will firstly be necessary to survey the personal attributes of driving licence applicants once more, as was already done in the past for the equivalence studies relating to the TDT. At the same time, the described dynamic processes in further development of the theoretical (and practical) driving test underline ever more clearly the necessity to formulate common interim objectives for driving school training and driving licence testing as a basis for coming research and development work in the field of media-supported hazard avoidance training. The focus of this work should be placed on appropriate didactic strategies to convey and test competences relating to traffic perception and hazard avoidance.

*Kristin Kaufmann & Annika Dressler*

# Systematisation of situational traffic demands for innovative test item formats

## 1.18 Overview

As a learning and test medium, computers offer a diversity of possibilities for simulation of the driving tasks which novice drivers face in real traffic in corresponding teaching/learning materials and test items. The essential benefits compared to driving in real traffic are to be seen in the risk-free learning or testing environment, the flexible creation of such materials with the aid of modern media technologies, with opportunity to depict a broad spectrum of traffic situations, and the extensive possibilities for systematic variation of the underlying demands. The visualisation software "VICOM", which was elaborated by the TÜV | DEKRA arge tp 21 working group to support the further optimisation of driving licence testing in Germany, can here be mentioned as a specific example of how traffic demands can already be transferred to the computer, and how simulated traffic situations can henceforth be developed in practically any chosen form.

Given the fact that real traffic is an eminently complex system and any specific traffic situation seems unique in its interactions with the concrete framework conditions, questions naturally arise as to the particular traffic situations which are to be visualised[41] for training and test purposes, and how the selection or discarding of certain situational demands can be founded. It must already be stated in advance, that conclusive answers cannot be expected in the present chapter, because the aforementioned questions refer not least to the (content) validity of the corresponding teaching/learning materials and test items. Appropriate methodical research and empirical studies are necessary to be able to answer such validity questions; such studies, however, are still outstanding. Instead, the initial objective is here to describe the different aspects which must be taken into account in the systematisation and selection of traffic situations and traffic demands for computer-based visualisation in the context of a future driving test.

The necessity for systematisation and limitation of the possibilities for situational visualisation already becomes clear from the mere attempt to simulate demand situations in which novice drivers are overtaxed and thus frequently involved in accidents. An enormous diversity of attributes is available for incorporation into the visualisations of such situations, for example specific aspects of the depicted driving manoeuvre, the types and numbers of other road users, the behaviour of those other road users, the visibility of hazard cues, the spatial context of the situation (e.g. urban, rural) and the weather conditions. The definition of accident-relevant situational demands is a prerequisite for scientifically founded design decisions. In view of the high complexity of real traffic situations, however, it is not a simple matter to identify those situational conditions which are the actual causes for (novice driver) accidents. Furthermore, it would probably be short-sighted to merely "replicate" a constellation of known physical situation factors without taking into account the associated cognitive demands which represent a challenge for the novice driver. It is rather the case that a meaningful balance must be found between situational

[41] The following thoughts refer solely to visual information acquisition when driving. Sight is by no means the only sense utilised by road users for perception and orientation in a given traffic situation, but it has been proven by Sivak (1996), among others, that the majority of information gathered during driving is acquired via the visual channel.

aspects from both areas, i.e. both the physical environmental conditions and the corresponding demands placed on novice drivers in respect of information processing. To support the further discussion, a number of conceptual foundations are presented in the following Chapter 5.2, alongside possible approaches to a scientifically founded selection of appropriate traffic situations. Looking ahead to the design process for computer-based innovative test item formats, finally, possibilities for the combination of selection strategies are discussed (Chapter 5.3).

## 1.19 Scientific approaches to the description and selection of traffic situations

### 1.19.1 The notion of "situation" as a framework concept for the description of traffic environments

The term "situation", which is widely used in everyday language to describe a person's or object's "state", "position" or "circumstances", acquires domain-specific accentuation when used in various scientific contexts (e.g. sociology, traffic psychology, geography). In sociology, for example, reference is made to aspects of interaction and the associated diversity of perspectives, which are then characterised more specifically as "ego" and "alter" perspectives. The purpose of this distinction is not to describe the individual persons as such, but rather to unambiguously locate a possible view of the world, for example by ascribing motives, interests, intentions and objectives (Luhmann, 1981). In this way, it is possible to draw a simplified picture of reality, while at the same time emphasising the references and ego/alter interactions (Luhmann, 2012). If the focus is placed on an individual perspective of a particular road user, then "situation" – in accordance with the descriptive approaches of (traffic) psychology – is to be understood to refer to the integrated system of environmental conditions with which a person interacts. This interaction can be characterised uniquely by a set of information, knowledge and response options (Pew, 2008).

With regard to road traffic, Reichart (2001) defines a "traffic situation" as a given spatial and temporal constellation of those perceivable and not yet perceivable traffic-related variables from the working environment of a road user which could influence his future driving behaviour. In this context, a traffic situation is to be subdivided into a "driving situation" – a fraction of the traffic situation which is generally and thus objectively perceivable from the ego perspective ("visually perceivable traffic space") – and a "driver situation"; the latter describes the situation as it is actually perceived by the driver. This distinction is relevant from the point of view of cognitive psychology, because the driver's perception is highly selective: Of the countless pieces of information with which a person is confronted in natural situations, he will only be consciously aware of the small proportion which is processed attentively (O'Regan & Noë, 2001; see also the discussion on "bottom-up" and "top-down" perception processes in Chapter 2.2.1). The differentiation between "driver situation" and "driving situation" also seems expedient for the designing of computer-based visualisations: While the driver situation reflects the subjective conditions for perception, and thus inter-individual performance differences with regard to traffic perception and hazard avoidance, the underlying driving situations can be objectivised, described intersubjectively and thus designed systematically.

To be able to design and visualise traffic situations for presentation by computer, it is necessary that their constituent elements can be classified and described precisely. It is only through systematic arrangement of such elements that it becomes possible to develop demand situations which are mutually similar or qualitatively different (e.g. with regard to content parameters or the degree of difficulty). Given the aforementioned high complexity of the traffic environment in which a particular driver participates in motorised traffic, systematisation is an inherently demanding undertaking, in which very different levels of differentiation are possible.

At a relatively coarse level, in accordance with Bernotat and Käppler (1985), a meaningful distinction can be made between the “natural environment” (e.g. flora/fauna, topography, time of the day), the “built environment” (e.g. traffic network, road surfaces) and the “social environment” (e.g. road users, traffic density, interactions between involved persons). These three categories permit disjunct classification of the visually perceivable elements of “traffic spaces” or “traffic situations”[42], such that they at the same time reflect behaviour-relevant attributes of those elements. For elaboration of a further differentiated basis for the description of elements of natural, built and social environments relating to traffic spaces, it is expedient to refer to existing terminologies such as those which are anchored in the legal framework (Road Traffic Regulations, EU directive on driving licences) or else in use in driver training (Learner Driver Training Ordinance, teaching/learning media). Last but not least, the catalogue of driving tasks for the practical driving test (see Chapter 6) also refers to elements of the traffic space, for example in its detailed descriptions of driving demands in connection with “Crossroads and junctions”, “Rail-borne vehicles” or “Curves”.

In Fig. 30 below, relevant attributes of the natural, built and social environments are presented for the example of the traffic space “Curves”. The blue highlighting illustrates how specific traffic situations can be described through the selection of elements from the natural, built and social environments. The diagram also shows that – on the basis of the catalogue of driving tasks for an optimised practical driving test (Sturzbecher et al., 2014) – further traffic spaces can be described alongside the featured space “Curves” (e.g. roundabouts, crossroads/junctions). Each traffic space can thus be assigned specific design attributes, the elaboration and combination of which enable the construction of a specific traffic situation.

[42] It is suggested that the term “traffic space” be used to describe separate traffic environments in the sense of fundamentally distinct infrastructures (e.g. curves, junctions). A “traffic situation” arises when a superordinate traffic space is defined in further detail in respect of specific attributes (e.g. topography, time of the day, buildings, actions of other road users).

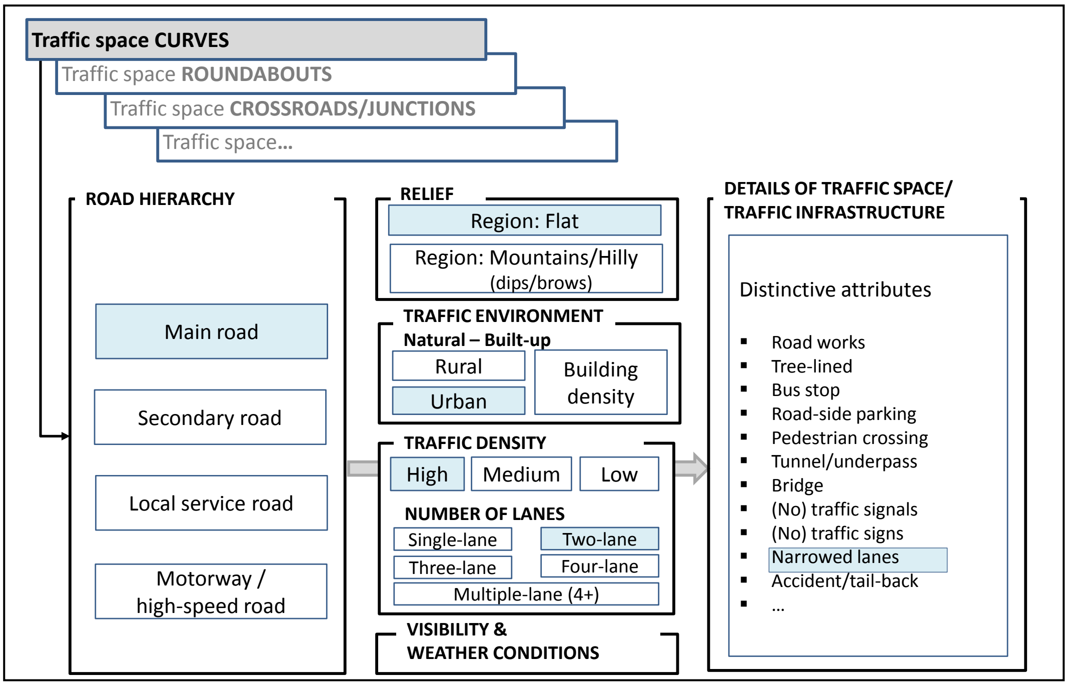



**Fig. 30: Example for the systematisation of elements of a "traffic space"**

With regard to the initially formulated necessity to define situation attributes which are especially significant for the frequency of accidents in real traffic, it must now be asked whether there are traffic spaces or traffic situations which entail a particularly high accident risk for novice drivers. The challenge to be solved when elaborating suitable visualisations for innovative test item formats is – from the multitude of possibilities – to identify those combinations of traffic infrastructure attributes which play an important role in accidents involving novice drivers, and to set tasks where abilities relating to traffic perception and hazard avoidance are decisive for successful completion. The two following Chapters 5.2.2 and 5.2.3 refer to precisely this challenge.

## 1.19.2 Situation selection on the basis of critical accident-relevant situation attributes

The international scientific literature includes numerous studies in which novice driver accidents are described in detail with regard to specific situation attributes or the determined causes. Major causes of accidents identified by Clarke et al. (2005), for example, are rear-end shunts and other collisions with a preceding vehicle (according to Whelan et al., 2002, hazards in the driver's own lane, in particular, are neglected), loss of control when negotiating bends, and accidents in darkness. Male novice drivers are more frequently prepared to take risks compared to female novice drivers, and exceeding of the permissible speed limit is a main, situation-independent cause of accidents. The lack of driving experience (Schlag et al., 1986), the associated failure to observe adequate safe distances and a correspondingly limited ability to react appropriately to unforeseen circumstances, on the other hand, are important accident causes for both sexes. Specifically, this could mean incorrect decisions between braking and swerving to avoid an obstacle, or else incorrect judgement of the safe stopping distance (McGwin & Brown, 1999). Both aspects contribute to accidents.

Single-vehicle accidents, where the driver has lost control of his vehicle, are by far the most common type of accident involving novice drivers. In most cases, the loss of control is a result of self-overestimation on the part of the driver, underestimation of a hazard or distraction. At the same time, young novice drivers are also over-represented among the persons responsible for frontal collisions or accidents at intersections (Clarke et al., 2006 for Great Britain; Laapotti & Keskinen, 1998 for Finland; Harrison et al., 1999 for Australia; McKnight & McKnight, 2003 for the USA).

For Germany, Grattenthaler et al. (2009) provide a overview of the most frequent types and causes of accidents involving drivers aged 18 to 24 years after analysis of the annual accident data published by the Federal Statistical Office (see Fig. 31).

| | Year 2003 (Federal Statistical Office, 2004) | | Year 2004/Year 2005 (Federal Statistical Office, 2005; Federal Statistical Office, 2006) | | Year 2006 (Federal Statistical Office, 2007) | |
|---|---|---|---|---|---|---|
| **Two most frequent accident types[1]** | Single-vehicle accident[3]<br>Accident in longitudinal traffic[4] | | Single-vehicle accident<br>Accident in longitudinal traffic | | Single-vehicle accident<br>Accident in longitudinal traffic | |
| **Five most common causes of accidents[2]** | Inappropriate speed | | Inappropriate speed | | Inappropriate speed | |
| | Safe distance | | Right of way | Safe distance | Safe distance | |
| | Right of way | | Safe distance | Right of way | Right of way | |
| | Turning at junctions | Driving in an unfit state | Turning at junctions | Driving in an unfit state | Turning at junctions | Driving in an unfit state |
| | Driving in an unfit state | Turning at junctions | Driving in an unfit state | Turning at junctions | Driving in an unfit state | Turning at junctions |
| **Age group** | 18-20 years | 21-24 years | 18-20 years | 21-24 years | 18-20 years | 21-24 years |
| Data source | Accident figures for 2003 | | Accident figures for 2004 and 2005 | | Accident figures for 2006 | |
| | Road accident statistics, based on police accident statistics | | | | | |

1 Accident type: Principal classification of accidents with injury to persons or fatalities
2 Accident causes: Incorrect driver behaviour resulting in an accident with injury to persons
3 Single-vehicle accidents are accidents where the driver loses control of his vehicle without other persons or vehicles being involved
4 Accidents in longitudinal traffic are accidents arising from a conflict between road users travelling in the same or opposing directions

**Fig. 31: The most frequent types and causes of road accidents involving drivers aged 18 to 24 years in Germany, in descending order of frequency (Grattenthaler et al., 2009)**

According to these data, single-vehicle accidents were the most frequent accident type in both age groups in 2003, 2004/2005 and 2006, followed by accidents in longitudinal traffic. The five most common causes of accident were inappropriate speed, failure to observe an adequate safe distance, failure to observe right of way, incorrect turning at junctions, and driving in an unfit state. The ranking of accident causes established by Grattenthaler et al. (2009) is also mirrored for the most part in the accident statistics for 2012 (see Fig. 32).

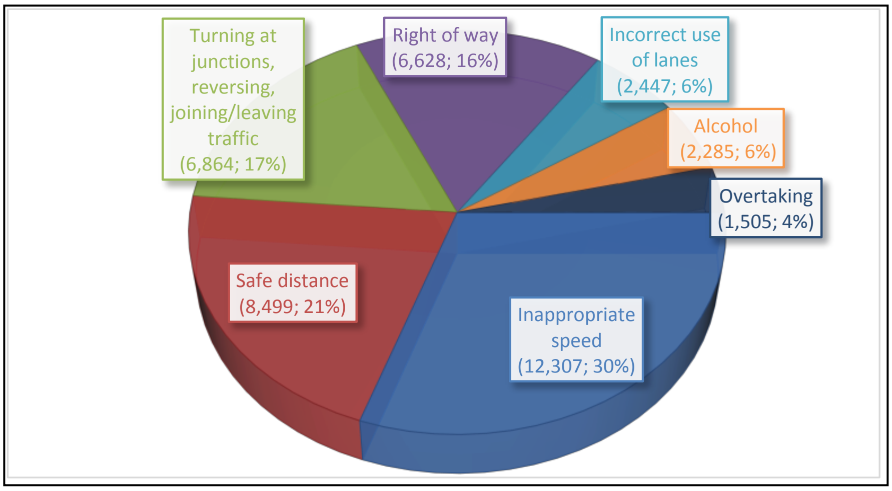



**Fig. 32: Incorrect behaviour of drivers aged 18 to 24 years in road accidents resulting in injury to persons (based on Statistisches Bundesamt, 2013)**

From the above diagram, which shows the aspects of incorrect behaviour which contributed to road accidents resulting in injury to persons for drivers in the age group from 18 to 24 years, it can be seen that driving at inappropriate speed (30%) and driving with inadequate safe distance (21%) together accounted for more than half of these accidents.

To achieve a differentiated analysis of accident causes, McKnight and McKnight (2003) classified road accidents involving more than 2,000 drivers aged between 16 and 19 years (see Fig. 33) with regard to the elements of incorrect behaviour which led to the accident. Combinations of several errors were also possible. As can be seen from the figures below, McKnight et al. (ibid.) identified deficits in traffic perception, i.e. both inattention and inadequate scanning of the driving environment (ahead, to the side, to the rear), poor speed adjustment and the failure to observe proper safe distances and clearances as significant accident causes. Neale et al. (2005) also name the inappropriate focussing of attention (due to fatigue, concentration on "non-critical" aspects of the driving environment or distraction through secondary tasks) as a key factor contributing to accidents on the basis of a one-year study of the accident involvement of 100 vehicles in real traffic.

| Behaviour | % | Behaviour | % | Behaviour | % |
|---|---|---|---|---|---|
| **Basic control** | **8.0** | **Search ahead** | **19.1** | **Adjusting speed** | **20.8** |
| Lane keeping | 2.6 | Distance | 3.1 | Traffic/road conditions | 8.7 |
| Turning path | 1.3 | Roadsides | 4.3 | Curves | 6.1 |
| Braking | 1.3 | Before left turns | 4.8 | Slick surface | 2.3 |
| Turning speed | 0.7 | Car ahead | 3.1 | Slick curves | 1.5 |
| Other | 2.1 | Left-turning vehicle | 2.9 | High speed | 0.7 |
| | | Next lane | 0.9 | Other | 1.5 |
| **Traffic controls** | **5.6** | **Search to the side** | **14.2** | **Maintaining space** | **9.8** |
| Traffic lights | 1.7 | Intersection: Burdened | 7.7 | Following distance | 5.8 |
| Stop signs | 1.3 | Intersection: Privileged | 5.5 | Crossing and entering | 1.4 |
| Lane use | 1.5 | Sight obstructed | 0.8 | Side clearance | 1.3 |
| Passing | 0.6 | Other | 0.2 | Overtaking | 1.1 |
| Other | 0.5 | | | Other | 0.2 |
| **Attention** | **23.0** | **Search to the rear** | **9.4** | **Signals** | **1.2** |
| Maintaining attention | 18.6 | Slowing | 3.0 | Interpreting signals | 0.8 |
| Avoiding distractions | 3.8 | Backing | 2.1 | Signalling intent | 0.3 |
| Attention sharing | 0.07 | Periodically | 2.1 | Signalling presence | 0.1 |
| | | Changing lanes | 1.5 | | |
| | | Other | 0.7 | | |
| **Driver-vehicle** | **6.3** | **Other search** | **0.9** | **Emergencies** | **9.4** |
| Alcohol impairment | 2.4 | | | Swerving | 5.6 |
| Fatigue | 1.7 | | | Skid recovery | 1.4 |
| Vehicle | 1.5 | | | Braking | 1.0 |
| Other | 0.7 | | | Tyre failure | 0.7 |
| | | | | Brake failure | 0.7 |

**Fig. 33: Percentages of road accidents attributable to specific incorrect behaviour on the part of young drivers (McKnight & McKnight, 2003)**

The systematisation approaches presented here provide a first overview of the current knowledge and the degree of detail which is available to describe accident-critical situation attributes. Both the international scientific literature and annual accident statistics for Germany contain concrete information relating to accident types and accident causes. Targeted and systematic utilisation of such information in the work to visualise demand situations for innovative test item formats could help to ensure that the demands presented to novice drivers at the computer are representative of those situations which pose difficulties for novice drivers in real traffic. Comparable selection strategies are already used in the Australian states of Victoria and New South Wales, for example, where the most common types of accident serve as a basis for the development of visualisations for use in both test items and teaching/learning materials (Genschow, Sturzbecher & Willmes-Lenz, 2013).

## 1.19.3 Situation selection based on the cognitive psychology relevance of hazard cues and their attributes

This chapter returns to the question of typical novice-specific competence deficits (see Chapter 2) and outlines, in particular, deficits in respect of visual perception processes. In the complex traffic environment from which decision- and action-relevant information must be acquired, these perception deficits must be placed against the necessity to recognise hazard cues. Such hazard cues are here first described in more detail on the basis of their specific attributes. On this basis, a suitable approach for the systematic visual representation of hazard cues at the computer is to be presented.

With regard to traffic perception, it can generally be assumed that novice drivers – compared to experienced drivers – will more frequently fail to gather a comprehensive picture of a given traffic situation (Benda & Hoyos, 1983). With this inadequate information base, they subsequently also find it more difficult to distinguish between hazardous and non-hazardous situations (McKenna & Crick, 1997; TÜV | DEKRA arge tp 21, 2013). Such false judgements of the situation may then lead to action errors, which in turn increases the probability of accident.

Underwood and Crundall (1998) demonstrated that inefficient visual search strategies, above all under complex driving conditions, are an important reason for incomplete information acquisition on the part of novice drivers and thus for their fragmentary or inappropriate representation of the environment (Deery, 1999). Despite the evidently different levels of hazard, novice drivers do not adapt their scanning techniques to different basic traffic situations, e.g. rural, suburban or multi-lane roads (Falkmer & Gregersen, 2001). Furthermore, they often make inadequate use of the side mirrors for information acquisition (Crundall et al., 2002). Novice drivers also recognise fewer hazard cues than experienced drivers (Whelan et al., 2002; Underwood et al., 2005; Borowsky et al., 2010). Spatially distant hazards, in particular, are perceived less successfully and in part judged incorrectly (Brown, 1982; Brown & Groeger, 1988). Novice drivers are moreover slower to discover visual hazard cues (Wallis & Horswill, 2007; Smith et al., 2009; Huestegge et al., 2010; McKenna & Crick, 1997) compared to experienced drivers; their reactions to perceived hazards are similarly slower (Scialfa et al., 2011, Deery, 1999, Crundall et al., 1999; Sagberg & Bjørnskau, 2006). The peripheral vision of novice drivers is less well trained to draw information from the traffic space than that of experienced drivers (Mourant & Rockwell, 1972) and, according to Soliday (1974), static elements are better perceived as hazards than dynamic elements.

As a prerequisite for the development of visualisations for computer-based use, in the context of the aforementioned information processing deficits of novice drivers, it is necessary to identify and characterise the visual elements of a traffic situation which influence hazard perception. This enables the systematic elaboration of those demand constellations which pose difficulties for novice drivers. With reference to the demands placed on schema-driven information processing, Crundall et al. (2012) distinguish three categories of visualised hazard cues:

a) Hazards which can be recognised through behaviour anticipation ("behavioural prediction hazards"): The source of a potential hazard is already visible before the hazard itself arises (e.g. a child is visible between parked cars, but only later steps onto the road).

b) Hazards which can be derived from environmental circumstances ("environmental prediction hazards"): The later source of a hazard is initially obscured by elements of the environment, but the context offers pointers to its existence (e.g. the limited visibility when approaching a tight bend is a cue to the potential hazard of an obstacle, e.g. a broken-down truck, blocking the road ahead).

c) Hazards which require attention to be divided between more than one precursor ("dividing and focussing attention scenarios"): Attention must be paid to at least two potential hazards, each of which could belong to either of the first two categories (e.g. if two pedestrians wave to each other across the road, it is probable that one of them will step onto the road).

Further hazard classifications in connection with hazard perception have been proposed by various authors. Chapman et al. (2002), for example, distinguish "chronic hazards" (e.g. driving fast) and "specific hazards" (e.g. a child on the road or a vehicle about to pull out of a parking space), which, in turn, can be further differentiated as "central hazards" (hazards in the focus of the driver's vision) or "peripheral hazards" (hazards at the edges of the driver's field of vision). According to Vlakveld (2014), furthermore, hazards are not necessarily "imminent threats", but may also arise as "overt latent hazards" (e.g. a visible road user who could behave in a hazardous manner) or "covert latent hazards" (e.g. a road user who could be on collision course, but is presently obscured). The latter category of hazards can be taken to correspond to example b) in the classification presented by Crundall et al. (2012) (see above).

Underwood et al. (2013) refer to the aspect of the suddenness with which a hazard cue arises and distinguish between the "abrupt onset" and "gradual onset" of a hazard. An example for an abruptly arising hazard could be a pedestrian who steps onto the road without prior warning (cf. Sagberg & Bjørnskau, 2006). Gradually arising hazards, on the other hand, demand anticipation of the developing situation on the part of the driver. In this case, conclusions as to the possible further development must be drawn from what is currently visible; the fact that children are playing at the roadside, for example, must be recognised as a cue to the possibility that one of them could run in front of the vehicle.

Various terminologies are used in the scientific literature to classify the visual perception of hazard cues. Nevertheless, the essence of the concepts presented here fits into the following overarching schema for description of the attributes of hazard cues (see Fig. 34), which is based on the three attribute dimensions "concreteness", "visibility" and "speed".

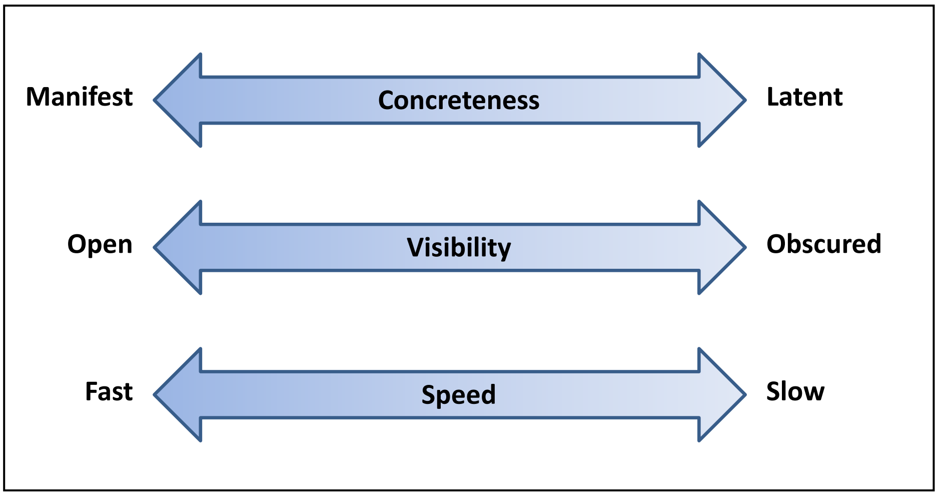



**Fig. 34: Schema for description of the attributes of hazard cues**

The cognitive tasks to be performed in connection with the recognition of hazard cues cannot be viewed separately from the specific attributes of the cues in question. A schema for description, such as that offered in Fig. 34, is able to support a systematic approach to the computer-based visualisation of traffic situations, by locating a certain relevant hazard cue (and, where required, varying this location) according to its particular attributes in each of the three aforementioned dimensions. Taking the example of a driving (driver) situation in which a vehicle is travelling along a road through a residential area, a pedestrian on the pavement would initially constitute a latent hazard cue. If this pedestrian moves towards the kerb in order to cross the road, however, his behaviour would represent an increasingly manifest hazard cue ("concreteness"). Parked vehicles along the side of the road could mean that the pedestrian is more or less obscured as he prepares to cross the road ("visibility"), and he could also be walking rather slowly or else very briskly towards the road ("speed").

It seems appropriate to apply such dimensional attribute descriptions, because the attributes of relevant hazard cues are naturally subject to constant change in a dynamically developing traffic situation. The demands to be met to complete a corresponding test item at the computer can be modulated through specific variation of the characteristics of the hazard cue to be detected in accordance with the defined dimensions. To this end, it is meaningful to refer also to existing results from empirical studies:

- For the dimension "concreteness", for example, it can be assumed that especially strong differences exist between novices and experienced drivers with regard to the anticipation of latent hazards (which must be predicted on the basis of the present environment), and that these hazards thus possess particular relevance for the accident involvement of novice drivers (Borowsky et al., 2009; Crundall et al., 2012; Sagberg & Bjørnskau, 2006).
- With regard to the dimension "visibility", empirical findings indicate that the demands placed on knowledge-based anticipation and schema-driven visual searches increase in line with the number and diversity of objects in a person's field of vision ("clutter", see Wickens & Hollands, 2000; Wickens & McCarley, 2007). The allocation of cognitive resources is also determined by the salience of hazard cues and irrelevant objects, i.e. the proportion of attention which they attract in the context of the overall stimulus configuration (e.g. due to their movement, colour, size or novelty; Itti & Koch, 2000; Wolfe & Horowitz, 2004).
- In the case of the dimension "speed", the time window for recognition of the decision- and action-relevant hazard cues plays a major role in dynamic visualisations. It is here possible to control the fundamental demands relating to fast and efficient information processing by varying the (relative) speeds of the depicted objects and the distances between them (e.g. as reference to an impending collision).

Independently of the aforementioned dimensions, the development of computer-based visualisations of hazard cues must also consider where such cues should be placed in the field of vision. Under the aspect of "information cost" (cf. SEEV model, Wickens & McCarley, 2007), for example, it can be expected that hazard cues are more difficult to detect, the farther they lie from the centre of the driver's field of vision (i.e. the most frequent area of fixation serving vehicle control). Such effects could be more pronounced among novices compared to experienced drivers due to their less-developed search schemata (e.g. where relevant hazard cues can only be recognised in the vehicle mirrors).

## 1.20 Combination of selection strategies as a starting point for computer-generated visualisations of traffic demands

As indicated in the previous chapters, the description of traffic situations – and the associated demands which must be met by drivers to master such situations – in accordance with accident-critical factors and cognitive psychology principles is a very challenging task. This is not only due to the fact that the traffic environment, as a "lifeworld domain"

(Sturzbecher, 2010), is an especially complex construct. It is equally the case that the available knowledge on accident-critical situation attributes is often insufficiently detailed as a source from which to derive concrete pointers for the computer-based visualisation of traffic demands. The secondary literature (e.g. annual accident statistics), for example, generally provides information on the frequency of accidents in certain traffic situations (e.g. when turning at junctions), but further demand-relevant situation attributes (e.g. time of the day) are not specified. It is precisely such more detailed knowledge regarding the frequency of accidents under certain combinations of attributes, however, which could supply important starting points for the visualisation of traffic situations. Databases with detailed information on accidents (such as that compiled by the GIDAS project[43]), which could be used to identify possible attribute patterns in the traffic situations concerned, are not freely accessible. This notwithstanding, it must be emphasised once more, however, that even the availability and consideration of more precise data on accident causes and accident-critical traffic situations are no substitute for empirical proof of the validity of computer-based question formats in teaching/learning and test media.

Which possibilities are available, therefore, to combine the aforementioned strategies and in this way to develop a practicable method for the elaboration of content-valid visualisations? At first sight, there appear to be two options, each of which has its own specific advantages and disadvantages. One approach would be to start out from infrastructural descriptions of traffic spaces, and then – as already illustrated with a concrete example in Fig. 30 – to combine different manifestations of the attributes which further characterise a particular traffic space. This would produce descriptions of specific, qualitatively distinguishable traffic situations, but the number of such descriptions would also be excessively high due to the countless possible permutations of attributes. The necessary limitation to a selection of relevant traffic situations could be achieved by assessing possible attribute combinations according to their significance for road safety. The computer-based visualisations could then concentrate on those traffic situations which, by way of suitable empirical methods (e.g. expert ratings), have been identified as high-risk situations with adequate reliability. Corresponding methodical approaches to description and risk assessment in the context of traffic situations can be found, for example, in McKnight and Adams (1971), von Benda (1985) and Fastenmeier (1995). The practical implementation of such methods, however, is expensive, as it is necessary to provide not only differentiated descriptions of traffic situations, but also assessment criteria for the risk assessment; at the same time, multiple assessments would be required to enable reliability evaluation.

An alternative approach could be based on empirical findings relating to novice-specific driving competence deficits (especially in connection with cognitive demands) and accident-critical situations. Such findings could be used to elaborate visualisations which (presumably) reflect the cognitive demands relevant to traffic perception and hazard avoidance. For illustration of such a method, reference can be made to the aforementioned study by McKnight and McKnight (2003). From their empirical study of accidents involving novice drivers, it can be inferred that inadequate visual hazard detection to the side is one of the most frequent causes of accidents, and that this applies in particular to the behaviour at crossroads and junctions. On the basis of these empirical findings, it

[43] The “German In-Depth Accident Study” is a joint project of the Federal Highway Research Institute (BASt) and the Research Association of Automotive Technology (FAT). The project began in 1999 and collects a broad diversity of data on accidents resulting in injury to persons, e.g. the circumstances of the accident, the behaviour of the persons involved, the injuries sustained, damage to the vehicles and the traffic infrastructure at the place of the accident.

would be possible to elaborate visualisations which contain the aforementioned situation attributes (crossroads) and the demands to be handled in this situation (visual hazard detection to the side). For modulation of the cognitive demands, furthermore, it is possible to refer to the elaborated attributes of hazard cues. Such an approach would be less expensive, because it is able to make use of secondary sources, but the implementation of visualisations would be dependent, above all, on the availability of suitable empirical studies. This, conversely, involves the danger that relevant aspects may be neglected when developing teaching/learning and test contents, simply because no reference is to be found in published research work.

It is evident that neither of the approaches outlined above is in itself satisfactory or sufficiently practicable as a method to develop computer-based visualisations for training and test purposes. Accordingly, it seems expedient to bring together the available knowledge as a foundation for the systematic selection or elaboration of suitable visualisations.

From the perspectives of pedagogical psychology and test didactics, the appropriate starting point would be the (still outstanding) closer definition of learning objectives pertaining to those traffic situations which are considered relevant and require corresponding competences to be assessed by way of a driving licence test. The cognitive schemata applied by experienced drivers to control attention focussing and early hazard detection, which were already described in Chapter 2.2.1, are one promising approach to the development of learning objectives. Research findings indicate that decisive importance must be attached to the traffic space, as the information cue pointing to the situation schema which is most appropriate. Borowsky et al. (2009), for example, demonstrated by way of a classification task that novice drivers tend to base judgements regarding the similarity of hazard situations on the actually depicted hazard cues (e.g. children). Experienced drivers, by contrast, referred additionally to characteristics of the traffic environment and classified the level of danger in a situation on the basis of the potential hazards which are more likely in the traffic space concerned. Correspondingly, Borowsky et al. (ibid.) conclude that prominence should be given to the close relationships between the traffic environment and the specific hazards which may arise in any given environment when designing teaching/learning media – or equally also test media – relating to traffic perception. In the Netherlands, a similar traffic-space-based schema theory is already being used as a starting point for the development of training programmes and test elements relating to hazard detection.

The above strategies relating to the systematic elaboration of visualisations could be addressed meaningfully in thoughts on schema-based competence acquisition and the derivation of corresponding learning objectives: For example, qualitatively distinct traffic spaces could at first be described by way of their infrastructural characteristics; some initial ideas on the structuring and description of traffic spaces in conjunction with the driving tasks of the practical driving test have already been presented above (see Chapter 5.2.1). Novice drivers could thus learn to recognise potential hazards – already at an early stage in their development – on the basis of a particular traffic environment ("traffic space") and its specific situation attributes ("traffic situation"). With regard to the diversity of traffic spaces, it would then be necessary to determine the specific situation attributes which are commonly recognised as cues to potential hazards by experienced drivers; these situation attributes should be contained in the visualisations of traffic situations and designed in accordance with present knowledge relating to the attributes of hazard cues (see Chapter 5.2.3). For judgement of the significance of different traffic spaces and traffic situations, and consequently of the learning objectives associated therewith, reference could be made to data from secondary literature (e.g. annual accident statistics), despite the fact that the available information is usually of a rather general nature (see Chapter 5.2.2).

The procedure outlined here is one example of how, through recourse to existing approaches to description of the traffic environment, as well as to scientific findings relating to novice-specific competence deficits and accident causes, it could be possible to visualise traffic situations which novice drivers apparently master less successfully than experienced drivers. The extent to which these visualisations actually meet such validity expectations can only be clarified by way of empirical studies using corresponding teaching/learning media and test item formats. With a systematic and properly founded approach to the visualisation of traffic situations, however, it seems by all means feasible to depict those traffic situations which are associated with novice drivers accidents in road traffic and to offer a risk-free opportunity to practice and test the appropriate behaviour by computer.

*Dietmar Sturzbecher, Susann Mörl, Tino Friedel & Mathias Rüdel*

# Evaluation and further development of the practical driving test

## 1.21 Overview

The present innovation report is the first to address work on the practical driving test (PDT) within the framework of reporting on further development of the scientific foundations of driving licence testing (see Chapter 1). It thus seems expedient to begin with a brief description of the fundamental objectives and potential of the PDT, and to provide an overview of the research and development work to date. The essential results of this research will then be outlined in the following Chapters 6.2 and 6.3, the former of which is devoted to studies conducted before the present report period 2011 to 2014. The concluding Chapter 6.4 then looks ahead to coming development steps and the benefits expected beyond the present report period.

The work on development of the PDT during the period covered by the current report (see Chapter 6.3) was essentially dominated by three projects. The first of these was the BASt project "Optimisation of the Practical Driving Test", which began in 2008 and ended during the first year of the report period (2011). The second project was a feasibility study conducted by the TÜV | DEKRA arge tp 21 working group between 2011 and 2012. The BASt project "Revision Project on an Optimised Practical Driving Test" ("Revision project"), finally, was commenced in 2013 and is scheduled to continue into 2015.

*Objectives and potential of the practical driving test*
From the perspective of psychological testing, the PDT can be categorised as a criterion-referenced or learning-objective-referenced test procedure (Hampel, 1977b). The content focus of the PDT rests on a test drive in real traffic, during which the performance of the driving licence applicant is assessed by way of systematic driving behaviour observation (Sturzbecher, Mörl & Kaltenbaek, 2014). The PDT serves to determine whether candidates have attained the specified training goals and thus possess the minimum level of driving competence necessary for safe and independent participation in motorised road traffic. In this sense, the PDT – like the TDT – fulfils both control and selection functions.

The control function of the PDT refers to the orientation it provides for the organisation of driving school training by specifying test contents, demand standards and assessment criteria: Both the training system and the individual driving licence applicants align their teaching and learning processes to the test contents, which, as common training standards, integrate the two aspects of driving school training and driving licence testing. The test results, in turn, represent feedback on the success of training and learning, and in this way point to potential for optimisation of the training effectiveness and further learning on the part of the test candidate. The PDT fulfils its selection function by ensuring that, in accordance with § 17 of the Driving Licence Regulations (Fahrerlaubnisverordnung, FeV), only those licence applicants who are able to demonstrate a safe, environment-aware and energy-saving manner of driving are entitled to participate in motorised road traffic. As is evidently the case, the driving competences which must be demonstrated during a test drive in real traffic, in particular, are for the most part precisely those competences which are imperative for later solo participation in road traffic. This founds the singular importance of the PDT within the overall system of novice driver preparation.

The overarching goal for the work on optimisation of the PDT is to improve its control and selection functions. To this end, it is a prerequisite that the content and methodology of the PDT test procedure are placed on scientific foundations, that it can be performed routinely, i.e. under defined conditions as an evaluation of more or less objective skills, and that it permits objective, reliable and valid determination of the status of the test candidate in terms of acquired driving competence (Lienert & Raatz, 1998). This requirement alludes to the acknowledged principal quality criteria for psychological testing (objectivity, reliability, validity), which must thus be satisfied by the PDT – alongside a series of secondary quality criteria such as economy, usefulness, reasonableness, resistance to falsification and fairness.

*Previous work on further development of the PDT*

In 2005, the Technical Examination Centres commissioned their joint development unit – the TÜV | DEKRA arge tp 21 working group – to force the pace of further development of the PDT, with the support of other scientific institutions, in order to attain the aforementioned overarching goal (see Fig. 35). The first outcome of these efforts was a research report published in 2008, in which the methodical foundations and possibilities for further development of the PDT were outlined as a starting point for subsequent research and development work (Sturzbecher, Bönninger & Rüdel, 2010). Building upon these results, the Federal Highway Research Institute (BASt) sponsored a follow-up project on optimisation of the PDT. Within the framework of this project, a comprehensive overview of the methodical foundations of the PDT was presented, and medium-term possibilities for optimisation of the system of novice driver preparation were elaborated (Sturzbecher, Mörl & Kaltenbaek, 2014).

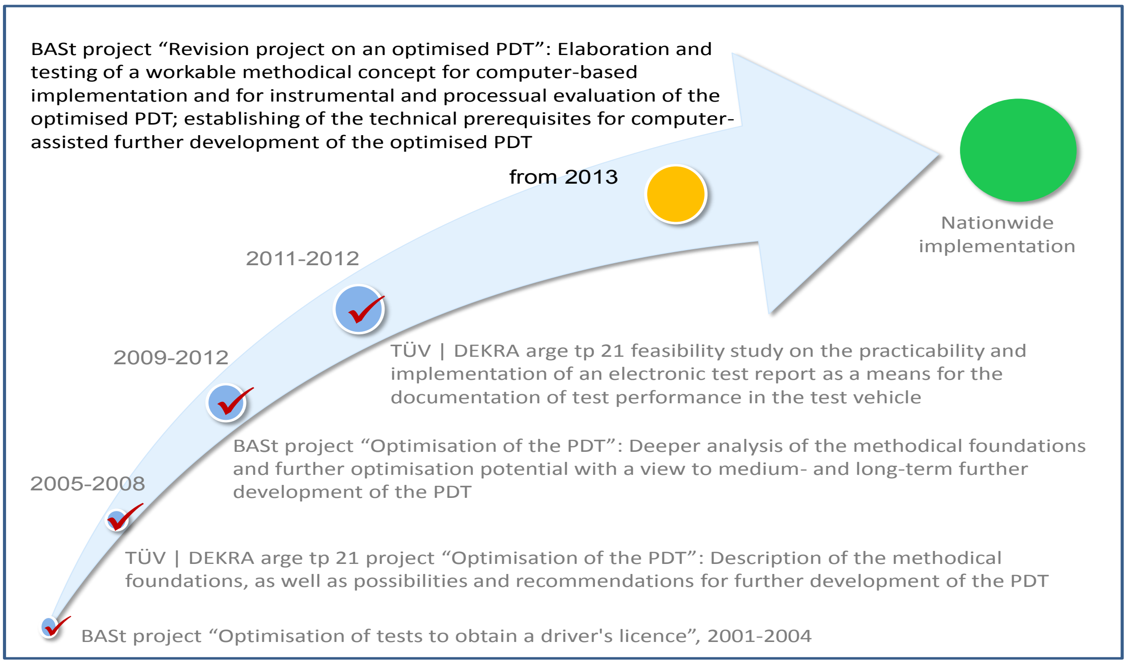



**Fig. 35: Research and development work on further development of the PDT**

One key aspect of the optimisation work is the development of an electronic test report to record and document candidate performance during the PDT. A corresponding feasibility study was conducted with a pilot implementation of the test report in 2011 and 2012[44]. The results and findings of these project stages, and similarly the contents of a revision project set up in 2013 to test all the elaborated methods and procedures of an optimised PDT, are described in the following chapters.

## 1.22 Foundations and starting points for further development of the practical driving test

Given that – as already mentioned above – the development of the PDT is here being discussed for the first time in an innovation report, the status of corresponding research and development work at the beginning of the report period is to be presented in slightly more detail in the following. The starting point is the project "Methodical foundations and possibilities for further development of the practical driving test" (Sturzbecher, Bönninger & Rüdel, 2010), which describes the theoretical, methodical and legal foundations of the PDT and maps their development since the 1970s. This fundamental description was supplemented with an international comparison. On this basis, conclusions and recommendations for further development of the PDT were described, and potential optimisation approaches were sketched.

In the course of this research project, the PDT was described from the methodical perspective as a competence-referenced diagnostic "work sample" (here specifically a driving sample), wherein the test performance of the driving licence applicant is observed, assessed and documented by a driving test examiner by way of "systematic (driving) behaviour observation": On the basis of the observed test performance, the examiner decides whether the test candidate possesses a certain minimum level of driving competence. This process of driving competence assessment was described in a circular model as an "adaptive test strategy". The adaptive test strategy comprises five action elements, which the driving test examiner realises more or less frequently, and in part also simultaneously, during the course of a test (see Fig. 36):

1. Planning and structuring of the test or observation situations by way of prescribed demand standards (driving tasks) and on the basis of knowledge of the test route, taking into account the performance already displayed by the test candidate, where appropriate
2. Systematic (targeted) observation of the candidate's behaviour in accordance with specified observation categories which define the aspects of behaviour and candidate competences to be examined
3. Assessment of the candidate's behaviour against specified assessment criteria, documentation of the assessment results by way of a test report, and elaboration of decision preferences for the final test result (pass or fail)

[44] This feasibility study was realised by the four Technical Examination Centres mandated to conduct driving tests (TÜV Rheinland Kraftfahrt GmbH, TÜV SÜD Auto Service GmbH, TÜV NORD Mobilität GmbH & Co. KG and DEKRA Automobil GmbH) and the Central Military Vehicle Registration Office of the Bundeswehr, under the aegis of the TÜV | DEKRA arge tp 21 working group and with support from the Institute for Prevention and Road Safety (IPV GmbH). The study was accompanied by a working group comprising representatives of the federal and state ministries responsible for traffic, the Federal Highway Research Institute (BASt), the Technical Examination Centres, the Bundeswehr, the driving instructors and other scientists working in the field, whose task was to discuss organisational, technical, economic and political questions relating to the electronic test report. The focus of the feasibility study was placed on the practicability of electronic documentation and the means for its implementation in the test vehicle (Friedel, Mörl & Rüdel, 2012).

4. Verification (reflection) of the current assessment and decision basis with regard to the attained degree of certainty and its suitability as justification for a valid decision
5. Decision-making relating to the planning of further observation situations and formulation of a final test decision.

The described adaptive test strategy of the PDT permits the driving test examiner to validate performance assessments reached over the (prior) course of the test: If the candidate completes a set task correctly, then he will not – as in a classic adaptive test – be confronted with progressively more difficult driving tasks as a basis for iterative localisation of the precise maximum of his relevant driving competence. It is rather the case that, in response to ambivalent test performances which do not permit an unambiguous assessment, a further, comparable driving task is set with the aim of minimising assessment doubts (Sturzbecher, 2010). Assuming consistent application of an adaptive test strategy, the PDT can be viewed as a partially standardised, criterion-driven test procedure, in which the results of planning, observation, assessment, verification and decision processes are combined purposefully to reach a reliable and valid judgement on the candidate's driving competence.

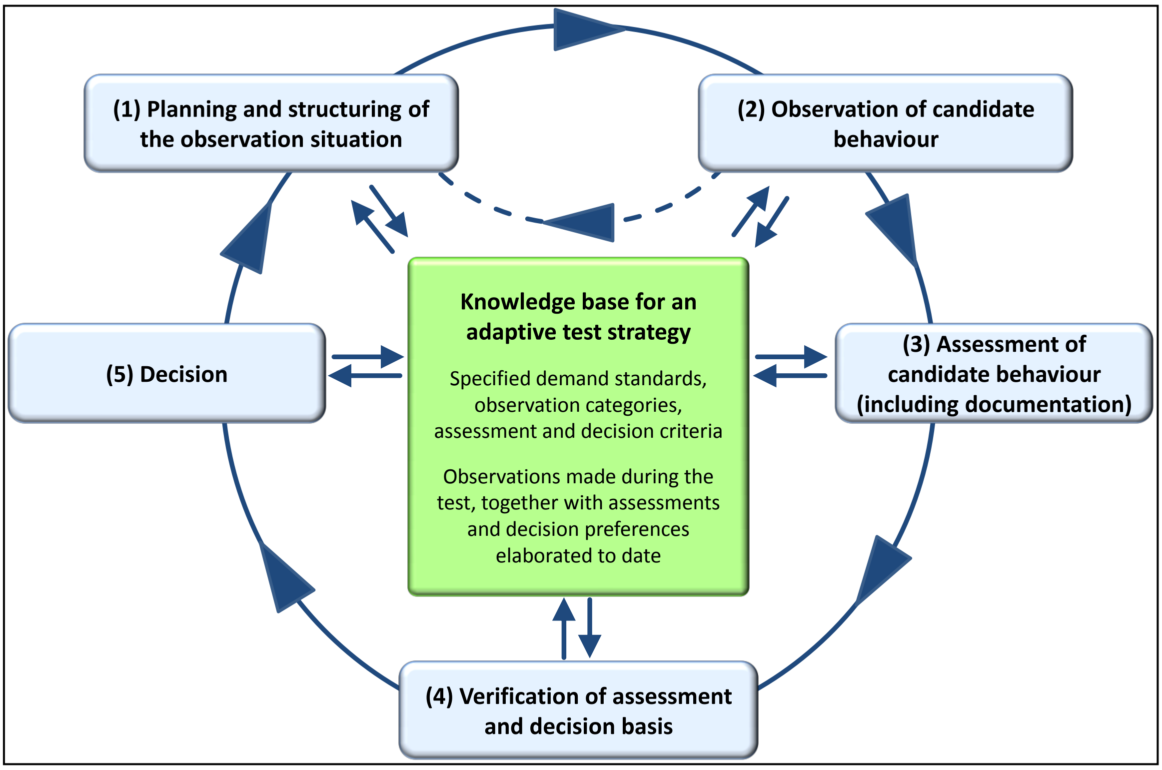



**Fig. 36: The adaptive test strategy (Sturzbecher, Bönninger & Rüdel, 2010)**

The adaptive test strategy represents the methodical control concept for an optimised PDT; it provides for systematic integration of the four elements of its process architecture and takes into account the special circumstances of test realisation in the lifeworld domain "road traffic", where planning and control are only feasible to a limited degree. The four elements refer to demand standards in the sense of (1) situation-related driving tasks (e.g. overtaking) and (2) situation-independent behaviour demands and dimensions of driving competence (e.g. traffic observation), as well as (3) assessment criteria and (4) decision criteria. In cooperation with further experts from the field, Sturzbecher, Bönninger und Rüdel (2010) presented a series of scientifically founded proposals for methodical optimisation of the aforementioned process elements and the implementation standards of the PDT. These proposals were elaborated further during the current report period (2011 to 2014) within the framework of the BASt project "Optimisation of the practical driving test", which is presented in the next chapter.

# 1.23 Methodical further development of the practical driving test during the report period

## 1.23.1 BASt project "Optimisation of the practical driving test"[45]

As already expounded, one fundamental goal of the research and development work on the PDT is to enhance the quality of the observation method – an instance of "systematic behaviour observation" – used for the test drive in real traffic. For methodical reasons, the observation situation must be structured on the basis of adequately standardised demands (here the so-called "driving tasks"); this represents the first step towards a guarantee of objectivity. Furthermore, systematic behaviour observation is dependent on so-called "observation categories" (here dimensions of driving competence) serving to focus the perception of the responsible observer (here the driving test examiner) on specific aspects of the behaviour displayed by the person under observation (here the driving licence applicant). In addition, it is necessary to provide assessment criteria by which to evaluate the observations. The observation categories and assessment criteria must also be adequately standardised; this can be seen as the second step with a positive influence on objectivity. Finally, a set of decision criteria is required to enable the individual task assessments to be condensed and translated into a test decision; the standardisation of these decision criteria is the third factor contributing to objectivity (Sturzbecher, Bönninger & Rüdel, 2010).

How is it possible to raise the methodical quality of systematic behaviour observations – such as those constituting the PDT? A promising answer to this question was given by Kötter and Nordmann (1987): On the basis of a comparative meta-analysis of published research on observation methodology, they elaborated a planning and control process which is intended to secure the methodical quality of observation results where – as in the case of the PDT, which is conducted many thousands of times each year – comparable results are the objective for extensive observation series. This planning and control process comprises three stages, which together describe the proper designing and optimisation of an observation method. It embraces the methodical planning and critical assessment of (1) the conception and structuring of the observation situation, (2) documentation of the data acquired in the observation situation, and (3) the evaluation methodology.

Within the framework of the project "Optimisation of the PDT" (Sturzbecher, Mörl & Kaltenbaek, 2014), this design concept for systematic behaviour observations, which is already well proven in psychological testing, was applied to the PDT. In this context, it is possible to identify four methodical and content-related challenges which must be overcome for successful further development of the PDT:

- Firstly, adequate implementation and development of an optimised PDT requires the elaboration of a demand concept wherein the specification of particularly safety-relevant, prototypical observation or traffic situations (in the following "driving tasks") with sufficiently standardised behaviour demands (in the following "dimensions of driving competence") serves to structure and control the test proce-

[45] This BASt project was conducted by the Institute for Applied Research on Childhood, Youth and the Family (IFK) at the University of Potsdam, in cooperation with the TÜV | DEKRA arge tp 21 working group.

dure in such a manner, that it yields meaningful (i.e. objective, reliable and valid) information on the driving competence of the test candidate. To this end, situation-related and situation-independent demand standards must be concretised and provided with a scientific foundation.

- Secondly, a discriminating evaluation methodology is required to permit high-quality psychometric observation and test results to be gained from the optimised PDT. This refers to the precise elaboration and uniform application of appropriately content-referenced and methodically sound assessment and decision criteria.
- Thirdly, it is necessary to establish a methodical and technological concept for adequate, computer-assisted assessment and electronic documentation of the candidate's test performance in the observation situations (the technological concept is hereafter embodied by the “electronic test report”).
- Finally, the elaborated implementation, demand, assessment and documentation standards must be made a subject of continuous further development in accordance with a (formative and summative) evaluation system.

In awareness of these four requirements, the first step was to elaborate a driving competence model as a basis for more detailed determination of the dimensions of driving competence to be verified by way of an optimised PDT. The initial starting point was the model proposed by Donges (2009), which merges action-oriented dimensions relating to the content-specific demand levels of the driving process with levels of action control according to Rasmussen (1983). Further inspiration for the description of driving competence was drawn from the spiral (driving) competence acquisition model by Grattenthaler, Krüger and Schoch (2009). On the basis of these two models, the areas of competence to be conveyed in driver training (vehicle stabilisation, vehicle manoeuvring, navigation, and values or attitudes) were defined and the components of driving competence to be assessed by way of the optimised PDT (primarily competence relating to vehicle manoeuvring, see below) were identified. Finally, following Klieme et al. (2007), it was shown that the demands to be met by test candidates must be described in the form of (minimum) training standards for novice driver preparation (Sturzbecher, Mörl & Kaltenbaek, 2014).

To further the development of training and test standards in the context of driving licence testing, the demand standards for an optimised PDT were described – on the basis of studies conducted by McKnight and Adams (1970a, 1970b) and McKnight and Hundt (1971) – in the form of situation-related driving tasks and situation-independent observation categories or dimensions of competence. Alongside, corresponding criteria for event-oriented performance assessment and overall competence evaluation were specified. In addition to the scientific foundations from the perspectives of educational and test psychology, the concretised demand and assessment standards for the optimised practical driving test also take into account the stipulations of the Third EU Directive on Driving Licences (Directive 2006/126/EC of the European Parliament and of the Council of 20th December 2006) and the national provisions of the Driving Licence Regulations (including the associated Examination Guidelines), as well as international experience and implementation practice relating to driving licence testing (ibid.).

The results of evaluation from the point of view of competence theory, analyses of the demands of driving a motor vehicle in public traffic in accordance with action theory, and the methodical construction of the PDT as an instance of systematic driving behaviour observation led finally to the formulation of a draft for a catalogue of driving tasks. This task catalogue describes the demand standards for the PDT. It contains first a total of eight situation-related driving tasks, some of which are further divided into subtasks: (1) Joining/leaving traffic and changing lanes, (2) Negotiating curves, (3) Passing and overtaking, (4) Crossroads and junctions, (5) Roundabouts, (6) Trams and railway level crossings, (7) Bus/tram stops, pedestrians and cyclists, and (8) Driving in a straight line. Secondly, the situation-independent observation categories or dimensions of competence to be assessed were likewise specified in the task catalogue and referenced to the aforementioned driving tasks: (1) Traffic observation, (2) Vehicle positioning, (3) Speed adaptation, (4) Communication, and (5) Vehicle control/Environment-aware driving. Thirdly and finally, event-oriented and competence-oriented assessment criteria were described for each driving task and each observation category: The event-oriented criteria define simple and serious errors, as well as examples of above-average performance; the competence-oriented criteria permit assessment of the candidate's performance (driving tasks and dimensions of competence) over the test drive as a whole on a four-level rating scale ("Very good", "Good", "Sufficient", "Inadequate"). The described catalogue of driving tasks also provides a content-related and methodical basis for the programming and testing of an electronic test report as a new means of test documentation. To support the necessary elaboration of documentation standards, recommendations were formulated with regard to hardware selection, ergonomic report design and the realisation of a feasibility study (ibid.).

Electronic test documentation is moreover a basis for future scientific evaluation of the PDT, and the theoretical concept for a corresponding evaluation system was described within the framework of the project; it is to comprise four complementary methodical elements: While the element "Instrumental evaluation" targets the psychometric quality of the methods employed by the optimised PDT, the objective of the remaining elements "Customer surveys", "Product audits" and "Analysis of test results" is to analyse the implementation quality of the test in regular use. The latter evaluation elements thus serve a processual evaluation and are intended to provide methodically sound proof of a uniformly high quality of test design and performance assessment across the whole country; as forms of external process evaluation, they correspond to the procedures of internal corporate quality management.

Against the background of the dynamic developments in vehicle technologies over the past few years, the project also offered professional recommendations for the fundamental treatment of driver assistance and accident avoidance systems in connection with realisation and assessment of the PDT. To this end, the functions and function principles of selected driver assistance systems were described as a basis for subsequent analyses of the possible influences of such systems on the acquisition and testing of driving competence. By way of examples, it was shown how the use of such systems may affect the performance of certain driving tasks and the assessment of corresponding dimensions of competence (observation categories).

In a draft for a "System Manual on Driver Licensing (Practical Test)", finally, a content-related and methodical concept was elaborated as a basis for implementation, continuous maintenance, quality assurance and further development of the optimised PDT. Alongside the institutional structures necessary to implement the envisaged system of testing, this draft described the planned methods and procedures – including the requisite demand, assessment, documentation and evaluation standards.

### 1.23.2 Feasibility study on the electronic test report

With regard to the necessary implementation of documentation standards for the PDT, the project described above suggested that an electronic test report must be introduced to enable the electronic recording and documentation of test performances. Such an instrument is deemed necessary to enable the driving test examiner to document the assessments reached in accordance with the defined assessment criteria in a proper manner and with acceptable outlay. Furthermore, an electronic test report is especially beneficial for the implementation of an adaptive test strategy, as the basic report layout offers the driving test examiner a constant overview of the driving tasks which have already been performed during the current test, together with the assessments recorded in each case; it also makes it easier to weigh up a final test decision. At the same time, continuous evaluation of the – anonymised – test results (in similar fashion to that implemented for the TDT) only becomes possible through the electronic recording and documentation of test performances. Last but not least, the provision of competence-referenced verbal and written feedback to the candidate is dependent on an adequately simple means to summarise the performance assessments which have previously been recorded electronically. The electronic test report thus plays a central role for an optimised PDT: It presents the content framework of the test to the examiner as a basis for the documentation process, and generates important information for performance feedback to the candidate and for optimisation of the test itself from the examiner's inputs.

Within the framework of a feasibility study on the electronic test report (Friedel, Mörl & Rüdel, 2012), the revised methodical system of the PDT – the restructured and streamlined driving tasks, the observation categories and the assessment and decision criteria (Sturzbecher, Mörl & Kaltenbaek, 2014) – was first programmed into corresponding application software. The resulting prototype of the electronic test report was then tested in practical use. The objective of these initial trials was not merely to determine whether the prototype of the program satisfied the defined requirements in terms of functionality and hardware or software design; equally important was the fundamental question as to whether the electronic documentation of test performance is actually practicable in the test vehicle, and if so, the extent to which it may be able to assist the driving test examiner in his daily work.

The feasibility study was divided into two phases: (1) A preparatory phase to clarify user expectations and to establish all necessary prerequisites for data collection, and (2) an implementation phase, during which the first empirical data on the handling of the electronic test report were collected and analysed.

The preparatory phase commenced in May 2011 with an analysis of the expectations of potential users of the electronic test report in the individual Technical Examination Centres and among selected driving test examiners. In this same context, work began on elaboration of a layout for the electronic test report (initially in the form of PowerPoint simulations). After planning of the study methodology, selection and procurement of the hardware, programming of the prototype of the electronic test report and training of the driving test examiners who were to take part in the practical trials, the preparatory phase was concluded in mid-September 2011. Already in conjunction with the aforementioned programming work, the first internal tests were conducted, and the original designs were adapted accordingly in the course of a series of development cycles and "feedback loops" (so-called "alpha testing").

The subsequent implementation phase lasted until February 2012; during that time, the electronic test report was tested first in 350 simulated tests and then in 600 real tests for a class B driving licence. An iterative process of optimisation was followed: The driving test examiners were contacted at regular intervals and questioned regarding their experience with the electronic test report by way of semi-standardised telephone interviews; this feedback was then taken into account in corresponding program updates.

As an outcome of the feasibility study on the electronic test report, it can be deemed proven that the content-related and methodical foundations of the initially presented concept for an optimised PDT (Sturzbecher, Mörl & Kaltenbaek, 2014) offer a workable basis for elaboration of a methodical and technological concept for the electronic (digital) test report. The software developed to implement the electronic test report, and likewise the chosen hardware solution, assist the driving test examiner in his planning of the further course of a test drive, and promote the transparent and objective realisation of testing (including test assessment, test documentation and the test decision); the electronic test report was considered a practicable instrument by a majority of the driving test examiners involved (see Fig. 37).

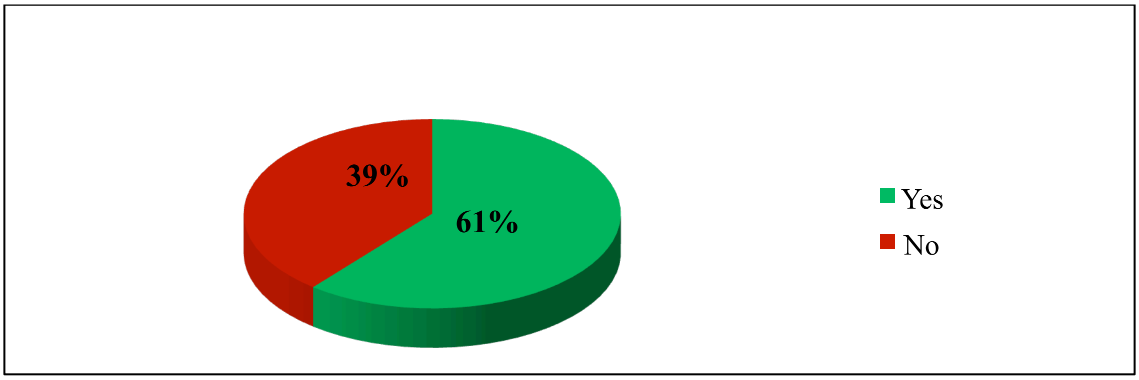



**Fig. 37: Assessed practicability of the electronic test report in the PDT**

As was to be expected, the practical trials revealed a number of points where further revision was considered necessary to improve handling of the electronic test report; for the most part, the corresponding modifications were already implemented within the framework of the feasibility study. It was shown furthermore that the prototype of the electronic test report must be optimised further in respect of the documentation of event-oriented assessments, especially in complex or relatively uncommon traffic situations. After completion of the feasibility study, therefore, the TÜV | DEKRA arge tp 21 working group conducted a comprehensive review of the operating logic and layout of the electronic test report. Further development of the electronic test report in preparation for full-scale deployment is to be accomplished by the subsequent revision project (see below).

The German Federation of Driving Instructor Associations (BVF) supported the development and testing of the electronic test report. In this context, consideration was also given to possible use of the report contents and operating concept for purposes of learner assessment during driving school training, and here in particular for determination of whether the learner is ready to take the practical driving test. Corresponding use of an electronic documentation program adapted to the specific demands of the training setting seems both expedient and desirable as a means to link practical driving instruction in the driving school with the later practical driving test (see also the previous discussion of joint training standards). Consequently, the BVF tested the prototype test report used in the feasibility study within the framework of 40 training drives. The feedback from the driving instructors indicated that their expectations were generally satisfied by the electronic driving performance report: The essential driving-task-referenced training contents upon which to base learner assessments are presented in the electronic test report and the corresponding performance of the learner could be recorded without problems during the training drive. On the basis of this test result, adaptation and expansion of the methodical and technological concept for an electronic report to cover also the important training contents of practical driving instruction is now a promising option. Beyond the original use in an optimised PDT, it would then be possible to develop an independent instrument for learner assessment during driving school training and to integrate this new means of assessment into the corresponding electronic teaching/learning systems.

## 1.23.3 BASt revision project on an optimised practical driving test[46]

The purpose of the BASt revision project is to continue development, optimisation and testing of the contents, methods and procedures of the optimised PDT – as was similarly the case for the revision project on the theoretical driving test conducted in 2008 – in preparation for full-scale deployment (Luniak, Mörl & Friedel, 2014). Accordingly, the targeted outcome is a workable methodical concept both for computer-based implementation and for instrumental and processual evaluation of the redesigned test. The project serves furthermore to establish the technical prerequisites for computer-assisted further development of the optimised PDT. Practical testing of the relevant methods and processes represents the core of the project, and is to be realised over a period of three months by the Technical Examination Centres mandated to conduct driving licence testing in four model regions. The essential project contents can be further divided into the following subprojects:

- The catalogue of driving tasks which describes the situation-related demands to be handled by the driving licence applicant (driving tasks) and the corresponding assessment criteria (descriptions of positive and negative behaviour when performing a set driving task, with reference to situation-independent observation categories or dimensions of driving competence) is to be revised and developed further to cover all driving licence classes.
- Building upon the results of the aforementioned feasibility study, the electronic test report is to be developed further and made available as a trial version for all driving licence classes.
- It is a general aim that, on the basis of the assessments of driving performance documented in the electronic test report, all driving licence applicants are to receive professionally sound and differentiated feedback on their current level of driving competence. To this end, it is necessary to elaborate and test a feedback system which – supplementary to the usual discussion with the examiner at the end of the test – provides a written summary of the errors observed, any above-average aspects of performance, and the areas in which driving competence could be improved. In contrast to current practice, this feedback is in future to be handed out to all candidates, irrespective of whether they have passed or failed the test, because test candidates must still be viewed as inexperienced drivers at the time of the driving test and would thus benefit from development-oriented recommendations for individual further learning (e.g. within the framework of accompanied driving).
- The contents and methodology of the optimised PDT need to be addressed within the framework of the initial qualification and further training of driving test examiners. Consequently, concepts are to be elaborated to achieve a uniform level of familiarity with the new system across the whole country and to train the handling of the new electronic test report. The further training concept is to serve also as a basis for training of the driving test examiners who take part in the practical trials of the revision project.
- Elements of the methodology for future continuous evaluation of the optimised PDT are to be developed further and tested accordingly. To this end, various studies are planned: For instrumental evaluation, several analyses are to be performed

[46] This project was also conducted by the Institute for Applied Research on Childhood, Youth and the Family (IFK) at the University of Potsdam, in cooperation with the TÜV | DEKRA arge tp 21 working group.

to assess the validity, reliability and objectivity of the observation method from the psychological perspective (e.g. analyses of observer consensus). Through evaluation of the test data collected by way of the electronic test report, conclusions are to be drawn regarding the demand profiles of particular test locations and the typical assessment patterns of driving test examiners. Further expected findings relate to possible correlations between certain test conditions and test results, and the driving competence deficits which frequently lead to the candidate failing the PDT. Customer survey instruments are also to be tested, firstly by way of a survey addressing those test candidates who took part in the practical trials of the revision project, and secondly with a survey directed at driving school owners.

- Last but not least, the necessary IT architecture for the optimised PDT is to be developed (further) and tested. Alongside completion of the electronic test report and the feedback system, this refers also to the development of a central database (including the corresponding administration program for the storage of all test contents and anonymised test and evaluation results) and the creation of a comprehensive technical infrastructure within the Technical Examination Centres (process adaptation).
- By the end of the revision project, a practice-tested and ready-to-use “operating concept” should have evolved as a basis for the implementation and evaluation of the optimised PDT (including an implementation concept). This operating concept is also be anchored in a “System Manual on Driver Licensing (Practical Test)” for the institutions involved in driving licence testing, and in a psychological process manual as reference for the driving test examiners. To this end, the existing draft of a manual, which describes the contents, methods and processes of the optimised PDT, must be amended in accordance with the knowledge gained from the revision project.

## 1.24 Outlook

Which benefits can optimisation of the PDT be expected to bring for training and testing practice in the future? Through the more precise definition and specification of situation-related and situation-independent driving behaviour demands (driving tasks and observation categories or dimensions of driving competence, respectively), the test contents are sharpened and better emphasis can be placed on the key relevant topics in driving school training and the overall teaching/learning process of novice drivers: In future, driving instructors, candidates and driving test examiners will all be able to consult the catalogue of driving tasks, which was elaborately jointly and in accordance with scientific standards by driving instructors, driving test examiners and experts from the field of psychological testing, as a source of precise information – in the sense of training standards – on the essential driving competences which must be trained and the levels of competence which must be attained. The common information base not only enhances demand and assessment transparency for the test candidate, but also contributes to closer integration of the systems of driver training and testing. This, in turn, strengthens the control function of the PDT and supports the implementation of a demanding and uniform driving test for all candidates.

It is furthermore expected that the safety impact of the PDT will be improved, as novice-specific driving competence deficits and accident causes were taken into account when defining the driving tasks and assessment criteria. This should enhance the selection function of the test.

The systematised and expanded assessment criteria[47], with their specific references to both individual driving situations and overarching dimensions of driving competence, permit a more precise statement as to the candidate's level of driving competence and thus differentiated feedback on his test performance. In future, this feedback is to be communicated to all candidates in two ways – based on the electronic test report and irrespective of whether or not the test was passed: First of all, the driving test examiner is to offer descriptive hints for further learning during a brief discussion at the end of the test drive; later, the candidate is to receive more detailed written learning recommendations either online or in printed form. This should help to raise the learning effectiveness of the PDT, and could also promote its safety-enhancing impact through a positive influence on the subsequent development of driving experience.

For the driving test examiners, the electronic test report should in future assist the efficient planning and realisation of tests, as it will permit them to call up all relevant information (e.g. the driving tasks to be tested and those which have already been performed, including the assessments recorded) at any time and in a single overview. In addition, the electronic recording of test data will facilitate test administration in the Technical Examination Centres, and the necessary coordination between the legislator and the Technical Examination Centres could also be modernised and simplified – as was already the case for the optimised TDT – through use of a joint, computer-based authoring system.

Furthermore, the electronic documentation of test performance opens up new possibilities for quality assurance; there is for the first time an opportunity to realise professional, formative and summative evaluation of the PDT. Within the framework of continuous evaluation, it could in future be determined, for example, which driving tasks are tested at which test locations, and to what extent given test locations are suitable for the testing of particular dimensions of driving competence. Supplementary to external monitoring of the effectiveness of the test in respect of its prime objectives, as demanded by the legislator, a deeper analysis of test results (beyond mere consideration of the pass rates) could also offer driving test examiners valuable information on their individual test behaviour in the sense of self-evaluation, and thus a basis for further development of their own, professional competence.

Finally, there are also new opportunities for an improvement of road safety and a reduction of novice driver risk: Scientific foundations for the contents and methods of the practical driving test will in future enable continuous and empirically supported further development of the demand and assessment standards on the basis of evaluations of test results and other traffic research sources such as accident analyses. Within the framework of such output control, it can be recognised, for example, whether a conspicuous proportion of the candidates fail to perform certain driving tasks correctly; corresponding analysis results can then be used to optimise driver training. In this way, monitoring of the test standards contributes to improvement of the training standards; the safety impact of novice driver preparation can be judged empirically.

The currently on-going revision project (see above) will show whether the above expectations do actually materialise in practice. Corresponding results and conclusions with regard to implementation of the optimised PDT are to be a subject of the next innovation report, which will cover the period from 2015 to 2017.

[47] The spectrum of event-oriented assessment criteria will in future cover not only normal performance, examples of simple errors and serious errors, but also examples of above-average performance.

*Mathias Rüdel & Dietmar Sturzbecher*

# Relevance of technical progress in vehicle engineering for novice driver preparation

## 1.25 Overview

Various content-related dimensions can be used to describe developments in mobility, and in the present context specifically in novice driver preparation, whose principal elements are driving school training and driving licence testing. This refers not least to traffic-related technical, legal and pedagogical dimensions. If these three dimensions are considered more closely from today's perspective, it is revealed that – apart from the effects of an accelerated rate of technical progress and increasing mutual influences between the aforementioned dimensions – further development of driving licence testing is dependent on the successful mastering of entirely new challenges, the responsibility for which lies above all in the Technical Examination Centres.

The new challenges are first and foremost a consequence of technical progress, as a fundamental driving force behind developments in mobility and in novice driver preparation. Technical innovations are entering the market at ever shorter intervals, and they are being installed in vehicles in countless variations. Impressive examples are, on the one hand, driver assistance systems with networked vehicle safety functions, but at the same time also new drive concepts. The first – already avidly awaited – solutions for (partially) automated vehicle control, and similarly the implementation of new vehicle concepts, illustrate the growing potential of technical innovation. All such innovations bring changes in the demands placed on driving behaviour and thus exert a decisive influence on the acquisition, assessment and upholding of driving competence. This concerns the pedagogical-didactic dimension of mobility and novice driver preparation: The driving instructors and the Technical Examination Centres, for example, will in future – to a greater extent than ever before – be required to judge continuously whether technical innovations influence the training- and test-relevant demand and training standards, and, if so, to determine how this may be manifested. On this basis, it is then necessary to adapt both the training specifications which control driver training and the test specifications which govern test realisation and assessment, and to implement the results of this further development in training and test practice. The essence of the training and test specifications must also be anchored in the pertinent legal provisions (e.g. in the Learner Driver Training Ordinance and in the Driving Licence Regulations), so as to guarantee uniform implementation and to enable quality assurance. Accordingly, the amendment of the legal framework must follow the pace of technical progress and will require regular redefinition of the systematic foundation; this relates to the legal dimension of mobility.

A review of the more than 100 years of development in driving licence testing against the background of the aforementioned legal dimension reveals a continuous process of specification and differentiation of the test contents and methods in the relevant statutory provisions. The legislative foundations for driving licence testing are spread between different sets of rules and regulations. The “Road Traffic Act” (StVG), the “Driving Licence Regulations” (FeV) and the “Guidelines for the Examination of Applicants for a Licence to Drive Motor Vehicles” (Examination Guidelines), the latter two each with their corresponding annexes, in particular, may be viewed as formally independent, but their contents are nevertheless closely related to each other. In the Examination Guidelines, the test specifications are codified with a high level of concretisation: Many points are described very exactly. Given the diversity and dynamics of technical innovation, and with each innovation the need to define corresponding rules, however, it seems hardly feasible to cast the test specifications for driving licence testing in such concrete legal provisions in the future, not least because the elaboration and execution of corresponding laws, regulations and guidelines would hardly be able to stay abreast of technical developments.

It can be taken from the further development of the pedagogical-didactic dimension of novice driver preparation, that the system of novice driver preparation is in the meantime viewed as an educational institution. Within society, the functions of educational institutions lie in the promotion, categorisation and selection of individuals. They first offer the individual the training (promotion) which will enable him to take on certain roles in society (Parsons & Shils, 1951). On the basis of the specific training contents and the results obtained, the individual then acquires rights to the assignment of certain roles (categorisation); at the same time, it is ensured that certain roles can only be assumed by suitable candidates (selection). This all applies equally to novice driver preparation and its two formal components, namely driving school training and driving licence testing. Driving school training can be considered firstly an instrument of promotion, but also serves to categorise the future participants in motorised road traffic by tailoring the training to a particular driving licence class, while the driving test provides for selection of the driving licence applicants and finalises their categorisation.

It is only in the last few years, since institutional classification and description of novice driver preparation from the perspectives of education and educational sociology (Genschow, Sturzbecher & Willmes-Lenz, 2013), that it has become possible to tackle systematic further development and control of this complex educational institution. Adequate consideration of technical progress is here one of the most important outstanding challenges: The urgency of the topic is evident and calls for a concerted approach among all those involved in training and testing. Precisely in this content-related field, however, the contribution from the side of training, i.e. driving competence acquisition, cannot end with the completion of basic driving (school) training, because knowledge and skills pertaining to the (safe) handling of innovative technologies are built up continuously and parallel to technical developments and their implementation. The use of in-vehicle innovations thus necessitates lifelong further learning on the part of drivers; to an ever greater extent, driving school training must be supplemented and continued by way of demanding further training measures!

In retrospect, the 20th century will be remembered as the century of motor vehicles with combustion engines. When operating such a motor vehicle, the driver must realise a multitude of control actions to manoeuvre the vehicle and not least to guarantee the safety of all those participating in road traffic. The 21st century, on the other hand, will probably enter the annals of vehicle engineering history as an era of pluralisation in respect of drive technologies and vehicle concepts – the option of "electromobility", with its potential to spare fossil resources, is recognised as especially promising. There is no mistaking the growing diversity of new vehicles and product designs, and further advances in vehicle technologies are already on the horizon. Later, we will probably even describe the coming years as the century of automated driving. Technical installations, such as driver assistance systems, are relieving the driver of more and more vehicle control tasks; to an increasing extent, they are also contributing prominently to hazard detection and avoidance.

To be able to determine the implications of foreseeable technical progress for the further development of novice driver preparation, it is first necessary to predict the scope of coming technical innovations and the periods over which they are likely to be introduced. Various partially automated driving functions, for example, are already in use in vehicles today. The added value of driver assistance systems in terms of comfort and – more importantly – in terms of road safety can only be exploited to the full, if the driver is aware of both the potential and the limitations of the systems at his disposal, and if their proper handling is integrated into his routine driving behaviour. At the same time, the capabilities of highly and fully automated driving functions suggest that fundamental changes lie ahead with regard to the role of the driver and his driving behaviour. These changes must be taken into account in the further development of novice driver preparation; on the other hand, their true impact is currently difficult to assess precisely, because the vehicle manufacturers have announced that the introduction of highly and fully automated driving functions is not to be expected before 2030.

Already in the meantime, however, the development of new drive and vehicle concepts, and with it the increasing diversity of motor vehicles, will stretch the traditional system of motor vehicle classification – and correspondingly referenced driving licence classes – to its limits as a basis for the description of driving competence demands. In the following chapter, therefore, brief overviews of the current state of development and the expected further development processes are to be given for each of the four phenomena mentioned above:

(1) Growing use of driver assistance systems,

(2) Increasing diversity of motor vehicles,

(3) Introduction of new drive and vehicle concepts,

(4) Implementation of highly and fully automated driving functions.

On this basis, probable consequences for the further development of novice driver preparation are to be derived from the pedagogical, didactic and legal perspectives.

## 1.26 Technical progress and its consequences for the acquisition and assessment of driving competence

*Growing use of driver assistance systems*

Human error is today considered to be the most frequent cause of road accidents (Gerster, 2013). It is not least for this reason that vehicle safety functions (driver assistance and automated driving functions) are in the meantime standard features of modern motor vehicles. Batch-manufactured driver assistance systems represent the first step in the direction of automated driving. “Driver assistance system” is the term applied to supplementary electronic systems which are installed in motor vehicles and there offer various forms of support to the driver in his handling of driving tasks.

Technical aids associated with driver assistance systems have already been available for around 20 years (Winner & Weitzel, 2012). Currently, however, new innovations with assistance functions are being added at ever shorter intervals. They offer support on all levels of vehicle control – stabilisation, manoeuvring, navigation and secondary activities – and are able to relieve the driver of different partial tasks (König, 2012). The vehicle is nevertheless controlled “with human intervention” (so-called assisted or semi-autonomous driving). In other words, the driver must monitor the system permanently and must be ready to resume full control over the vehicle at any time (Bundesministerium für Verkehr und digitale Infrastruktur, 2014).

The individual driver assistance systems can be classified as either “safety systems” (which serve to enhance driver safety), “intervention systems” (which intervene directly in vehicle behaviour), “comfort systems” (e.g. cruise control) or “information systems” (which provide information to the driver on the condition of the vehicle or the surrounding traffic environment). Depending on its individual functions, the purpose of a driver assistance system is thus to inform, warn, recommend or intervene. It is common to all systems, however, that they are intended to offer the driver the required assistance without taking away his fundamental responsibilities in road traffic (Sturzbecher et al., 2014). This includes also the widening of the performance limits of human perception and support in hazardous situations (Winner et al., 2012). Possible implementations of the human-machine interaction include camera-, laser-, ultrasound- and radar-based systems providing functions such as lane-keeping support, speed adaptation or 360-degree monitoring. Adaptive cruise control systems, as well as traffic jam, emergency brake, lane change, lane-keeping and parking assistants, permit automated braking, lane changing, acceleration, obstacle detection and collision avoidance (for an overview, see Sturzbecher et al., 2014). Driver assistance systems enjoy great popularity (Continental Mobility Study 2013) and their use is certain to become even more widespread.

Which future challenges arise for novice driver preparation from the increasing prevalence of driver assistance systems? Driver assistance systems are able to support the driver and thus enhance driver safety in certain traffic situations; that is not true in all situations and under all conditions, however. This means that both drivers and driving licence applicants, as well as driving instructors and driving test examiners, must be familiar with the benefits and limitations of individual driver assistance systems. They must be aware, in particular, that any technical system may be subject to functional limitations in certain situations. As a result, the driver must be able to monitor and, if possible, control the functioning of all the systems in use. One of the undesirable effects in conjunction with safety-relevant driver assistance systems is that their accident prevention capabilities are sometimes overestimated by inexperienced users and they may thus convey a false sense of security (e.g. "We have no need to worry because the car has ABS.")

To be able to exclude the described risks, basic knowledge relating to driver assistance systems is indispensable: Drivers, and likewise driving licence applicants, must know which driver assistance systems are installed in their vehicle, how these driver assistance systems function, which traffic situations or driving tasks are easier to master with the aid of the installed systems, and where the inherent limits of a particular driver assistance system lie. Both driving school training and driving licence testing should thus inspire driving licence applicants to investigate the driver assistance systems installed in their vehicle and to learn how they are used. On this basis, they should then also be encouraged to actually make use of the available systems and to gain experience of their potential and performance limitations. This could at the same time help to overcome the still inadequate use of driver assistance systems; statistics indicate that, for cost reasons, the first vehicles owned by novice drivers are often lacking modern driver assistance systems (Krüger, 2010). On the other hand, novice drivers often hire better equipped vehicles, or else use such vehicles within the family. Consequently, driving licence applicants must not only acquire the aforementioned basic knowledge in the course of their theory classes and practical driving instruction, but must also learn how to recognise the various driver assistance systems in different types of vehicle.

Turning to driving licence testing, it must be required that test candidates know how to make proper use of the available driver assistance systems in accordance with applicable regulations, and furthermore how to control their relevant functions – insofar as this is actually possible, given that some driver assistance systems cannot be deactivated or operated "consciously". On the other hand, it is especially important to ensure that a test candidate is able to drive adequately safely without certain driver assistance systems. It thus also seems evident that driving instructors and driving test examiners must possess sound knowledge and skills with regard to the use or non-use of driver assistance systems, and that certain pedagogical-didactic and legal specifications should be in place to govern the use of driver assistance systems in the training and test settings; this is necessary to ensure a uniform system of novice driver preparation throughout Germany and not least to guarantee test equality.

One possible starting point for work to overcome the challenges outlined above is the catalogue of driving tasks for an optimised practical driving test (Sturzbecher, Mörl & Kaltenbaek, 2014). This catalogue of driving tasks describes demand standards and assessment criteria – in the sense of training standards – relating to (1) Traffic observation, (2) Vehicle positioning, (3) Speed adaptation, (4) Communication and (5) Vehicle control/environment-aware driving. These demand standards and assessment criteria are applicable to both driving school training and driving licence testing; at the same time, their contents point to important areas in which driver assistance systems could take over certain tasks. The catalogue of driving tasks is thus to be seen as an orientation aid or "framework" for the systematic determination of training and test specifications.

*Increasing diversity of motor vehicles*
Over the past decades, a system of unambiguously defined vehicle and driving licence classes has been developed and serves as a fundamental basis for driving school training and driving licence testing. In the meantime, however, the growing diversity of vehicle concepts and variants – it is here sufficient to consider new groups of motor-driven vehicle such as e-bikes, trikes, quads, Twizys or Segways – means that the existing regulations are fast approaching their limits as criteria for vehicle and driving licence classification. Even today, it is sometimes difficult to fit innovative vehicle concepts into one of the existing categories. And it is already foreseeable that the future will bring many more new types of vehicle which defy assignment to a traditional vehicle class and thus no longer align with the present matrix of driving licence classes. Aside from the present complications with regard to legal classification, it seems rather more problematic that all participants in road traffic will in future find it increasingly difficult to associate particular characteristics (e.g. a high rate of acceleration) with the unfamiliar appearances of individual vehicles and to draw appropriate, experience-based conclusions for their own traffic behaviour. Such traffic experience could be acquired through corresponding advanced driver training, though very few such offers exist to date.

The increasing diversity of vehicle variants is not only a challenge with regard to further training opportunities for drivers; it is equally important for novice driver preparation to react to this technical phenomenon. One fundamental question to be answered in this context is whether the traditional practice of organising driving school training and driving licence testing for a particular driving licence class is a meaningful approach for the future, i.e. whether the technical classifications of motor vehicles must necessarily be congruent with the competence-referenced classifications of driving licences. Perhaps it would be better to convey fundamental, class-independent aspects of driving competence, such as the appropriate driving behaviour in specific traffic situations or when handling certain driving tasks, in the form of "basic driver training", and then to assess the level of competence acquired by way of a "basic driving licence test". Subsequently, further courses of driving school training and additional driving licence tests could address special vehicle concepts, possibly in accordance with a new, graduated concept offering different paths to the granting of a driving licence. The supplementary training and testing could then be geared to specific knowledge and skills associated with the new vehicle concepts, e.g. applicable laws of physics, appropriate behaviour with regard to traffic perception and hazard avoidance, or practical handling skills.

In practical terms, the outlined reorganisation of the process to obtain a driving licence would require that, each time new vehicle variants are introduced, it must first be clarified whether any specific explicit knowledge (e.g. special knowledge relating to driving physics or appropriate behaviour in the context of traffic perception and hazard avoidance) or specific practical skills are necessary to be able to handle and control the vehicle concerned. In the interest of public road safety, such specific knowledge and skills must be taught and tested in addition to basic driving competence. Parallel to this, the test contents, test methods and forms of testing would need to be developed further by the Technical Examination Centres. Such further development, however, should not be limited to mere expansion of the test contents and applicable forms of testing; it would instead be desirable to elaborate innovative proposals for potentially graduated driving licence careers, which could also involve thoughts on accessibility (e.g. age requirements), the transitions between individual stages and possibilities for combination (Bönninger, 2014a; Glowalla & Jordan, 2014).

*Introduction of new drive and vehicle concepts*
Alternative drive concepts and fuels are not new, and can in fact look back over a more than 180-year history. The first practical electric motors date from 1830 and are thus older than the earliest mobile applications of combustion engines, which were derived from the Otto cycle engine presented in 1876 (Thomes et al., 2013). Nevertheless, the inevitably forthcoming transition from combustion engines to electric drives still represents a major technical and organisational challenge. At the present moment in time, most electric vehicles possess a range of well less than 150 km; they are suitable above all for short journeys and are thus especially practical in urban settings. Experts see particularly good potential for the wider use of electric drives in connection with on-demand mobility (e.g. car sharing, ride sharing), as well as in mini and supermini cars, electric bicycles and scooters. In addition, attention must also be paid to technological innovations with the objective of further raising the efficiency of conventional petrol and diesel combustion engines (Rath & Bozem, 2013).

The numerous variants of hybrid vehicles can be distinguished on the basis of the underlying technology (micro-, mild-, full- and plug-in hybrids, in the two powertrain configurations "serial hybrid" and "parallel hybrid"). The range of products implementing hybrid drive technologies is thus very diverse; it seems likely that applications will emerge in practically all vehicle classes – with the exception of the smallest mini cars (Schlick, 2013). The hybrid vehicle categories "plug-in hybrid" and "range extender" (incorporating a small combustion engine with a generator) are also suitable for long-distance travel. A plug-in hybrid with increased energy storage capacity can cover up to 50 km in electric drive mode before a combustion engine takes over. The range extender technology functions with a specially optimised combustion engine serving as a generator to produce electricity (Rath & Bozem, 2013). Vehicles with fuel cells or hydrogen-fuelled drive ("fuel cell hybrid vehicles") are also propelled electrically. The necessary energy is generated directly on board, however, and does not have to be supplied from an external source; an additional combustion engine is not absolutely necessary. Fuels cells permit a greatly extended range, and today's fuel cell vehicles are already designed to cover distances of up to 800 km. The first series-manufactured fuel cell cars in Europe are expected to be available for sale in mid-2015 (ZEIT ONLINE, 2014). It is assumed that the introduction of fuel cells will revolutionise the way a typical vehicle powertrain is designed (Ludwig Bölkow Systemtechnik GmbH, 2006).

Which new challenges could arise for novice driver preparation from the introduction of innovative drive technologies and vehicle concepts? One immediately conspicuous aspect is the potential impact on demands relating to the competence dimension "Vehicle control/environment-aware driving": Firstly, new vehicle concepts will entail changes in the way a vehicle is handled compared to traditional types of vehicle; at the same time, it can already be noted today, that even functionally identical vehicle components are becoming increasingly manufacturer-specific and must thus also be activated and operated in different manners. Further new demands on the level of vehicle control may include changed performance characteristics, differences in road holding or an unaccustomed braking response; this must be taken into account in conjunction with vehicle positioning and speed adaptation, two further important elements of driving competence. Finally, driving licence applicants must be familiar with the ecological implications of a decision to purchase and drive a vehicle based on an alternative drive concept, so as to enable them to behave with due environmental awareness and to adapt their driving style accordingly. It thus cannot be excluded that the communication and testing of technical knowledge and skills, which have been losing importance since the 1970s, must play a greater role in driving school training and driving licence testing in the future.

The arrival of vehicles with electric drives and other alternative design concepts will also place new demands on road users in respect of the partial driving competence “Traffic observation”, which is of particular relevance for hazard avoidance. Due to the absence of engine noise, for example, electric vehicles are almost silent up to a speed of around 30 km/h; it is only at higher speeds that tyre and wind noise permit perception of the vehicle. The reduced noise emissions may seem desirable from the ecological perspective, but they can also constitute an increased accident risk under certain traffic conditions, especially for road users with impaired vision or hearing – which includes a growing number of older road users. This risk stems from the fact that, especially in circumstances with limited visibility (e.g. before the brow of a hill or in fog), road users tend to rely not only on their visual perceptions, but also on their hearing. This is scarcely possible with electric vehicles, however, and both motorised and non-motorised road users must still be made aware of this new situation.

While expansion of the electric car fleet in Germany has to date fallen far short of the expectations of economic and transport politics, there has been a marked increase in the numbers of traffic participants using e-bikes and other similar means of leisure-time mobility. At first sight, there is little difference in appearance between an e-bike and a conventional muscle-powered bicycle, but the former achieve far higher rates of acceleration and permit travel at higher speeds. This applies also to the simplest form of e-bikes with a speed-limited auxiliary drive, which are classified as bicycles under present legislation and can thus already be used in public road traffic by children and youths under 14 years of age – without any proof of theoretical knowledge (e.g. of traffic rules) or practical handling skills. It seems probable, on the one hand, that lack of an obligation to undergo training could entail risks for public road safety. On the other hand, carefully planned rules to control access to corresponding mobility entitlements could also provide opportunities for an earlier, graduated entry into motorised mobility, by offering young people the possibility to use low-powered and speed-limited electric vehicles. Especially in rural regions with sparse local public transport services, this could expand the training and employment options for young people, in addition to the gains in terms of road safety.

*Implementation of highly and fully automated driving functions*

Some first examples demonstrate that highly and fully automated driving is already today a technically feasible proposition; nevertheless, it is rather unlikely that any mentionable number of vehicles with highly or – in particular – fully automated functionality will be seen on the roads before 2030 (see above). Fully automated driving means that vehicle control is accomplished without human intervention. At this level of automation, it is no longer necessary for the driver to monitor the vehicle systems, because any applicable system limitations are already recognised by the system itself, which is in all situations capable of returning the vehicle to a state of minimal risk (Bundesministerium für Verkehr und digitale Infrastruktur, 2014).

Fully automated driving requires the transmission and processing of considerable amounts of data (e.g. via car-to-car communication or car-to-infrastructure communication), and that in compliance with data protection stipulations. To date, however, it seems that no adequately workable solutions exist, neither in the field of information and data processing, nor with regard to data protection and security (Bönninger, 2014b). Similarly, the elaboration of a legislative framework for automated driving is yet to be tackled, not only in Germany, but also worldwide (e.g. through corresponding amendment of the Vienna Convention). The Federal Ministry of Transport and Digital Infrastructure (BMVI) has thus established a round table "Automated driving", at which experts from various disciplines are to discuss and describe the conditions and demands of automated driving, along with possible scenarios for its introduction. The only consensus which has been reached so far, is that the medium- to long-term introduction of automated driving could serve to improve traffic flows and enhance road safety.

Which challenges arise for the further development of novice driver preparation in connection with the use of vehicles with highly or fully automated driving functions? While it is true that the introduction of such vehicles would alter the typical role of the driver, the Vienna Convention requires (for the time being) that vehicle control remains in the hands of the driver. The latter must retain responsibility for his vehicle at all times, and must be able to override or deactivate on-board assistance systems – irrespective of whether such systems provide partial or full automation – in case of technical defects or in limit situations (Sturzbecher, Mörl & Kaltenbaek, 2014). The relevant legislation may well be changed in the future, but until such time, it remains an essential task for the driver to be prepared and able to intervene into the highly automated driving processes at any moment, especially in critical situations in which the vehicle returns the command over a particular driving task to the driver. The role change will thus be manifested in an increasing proportion of time being allocated to monitoring tasks relating to vehicle control, while the proportion spent on the execution of manoeuvring tasks will decrease; even so, the driver must still be in a position to perform all necessary manoeuvres without assistance – as has always been the case.

From today's perspective, it is still difficult to foretell the final outcome of the expected role change and how quickly it will be implemented. In line with the aforementioned predictions of the automobile industry, however, every second person questioned within the framework of the representative Continental Mobility Study (2013) in Germany believed that automated driving would belong to everyday life in 10 to 15 years. If so, this would have immediate and direct implications for the organisation of driving school training and driving licence testing. It would be imperative, for example, to expand the current training and test contents to accommodate the acquisition and testing of knowledge and skills relating to the necessity and actual performance of corresponding monitoring tasks. As applies similarly with regard to traffic perception and hazard avoidance, this includes the recognition of signs for disturbed or suboptimal execution of the driving tasks by the automated control system (i.e. targeted information searches and processing), the reaching of decisions on whether or not to intervene, and the selection and successful execution of suitable intervention actions, where appropriate.

The shift in the nature of driver actions from psychomotor vehicle control to cognitive monitoring of the vehicle systems, as a natural consequence of the outlined role changes and the expected partial or full automation of driving functions, could possibly be a source of further risks. Routine elements of vehicle control could gradually become alien to the drivers of fully automated vehicles if they are no longer in regular use, for example, with possibly negative implications for safety should any of the assistance systems fail. It is also conceivable that the elimination of physical control actions and the limitation to relatively monotonous system monitoring could lead to new forms of fatigue and inattentiveness on the part of drivers. Such phenomena would then need to be addressed in driving school training and driving licence testing under the content heading “Human behaviour as a risk factor”.

## 1.27 Demands on the further development of novice driver preparation

It is indisputable that innovations in vehicle engineering necessitate further development of the systems of novice driver preparation and further driver training – that is convincingly clear from the given examples. Neither the precise pace of future technical progress nor its longer-term impact on the acquisition, assessment and preservation of driving competence can be predicted with any reliability; on the basis of the current and foreseeable circumstances, however, it is already possible to commence planning and implementation of the next development steps in driving school training and driving licence testing. The specific task for the Technical Examination Centres in this context is to further develop the content-related and methodical test standards in line with the technical advances; furthermore, they must contribute to the shaping of a future traffic system, so as to safeguard and further enhance the attained levels of road safety. When tackling these tasks, it is of prime importance to ensure reconciliation of the technical, pedagogical-didactic and legal dimensions.

From today's perspective, technical progress throws up a multitude of unanswered questions with direct or indirect consequences for the further development of novice driver preparation. It already seems certain, for example, that the technical diversity of motor vehicles will render it increasingly difficult to uphold the present system of driving licence classes and to elaborate a manageable and meaningfully structured system of class-independent and class-specific minimum training standards for the acquisition and testing of driving competence. It must thus be asked how the overall vehicle fleet will change over the next 10 years, and which conclusions can be drawn from this knowledge for the further development of driving school training and driving licence testing.

Throughout their tradition-steeped history, which began with the founding of the first “Steam Boiler Inspection Association” in 1866, the Technical Examination Centres have served as dependable testers of technical products and the operators of such products. As responsible mediators between the interests of human society, technology and road traffic, they have also been furthering development of the contents and methods of driving licence testing for more than 100 years (Brauckmann, Hähnel & Mylius, 2006). This longstanding experience must now be put to use to overcome a completely new challenge: Today, innovation in vehicle engineering is accelerating at an unprecedented rate, and necessary knowledge relating to the (safe) handling of innovative vehicle technologies is gathered simultaneously with the introduction and use of new systems (Mead, 1971; Sturzbecher, Kammler, Weiße & Breitling, 2009). To avoid a situation in which the users of technical innovations are only able to learn from mistakes, it is necessary to look into the future and to anticipate the demands which will one day arise for users of new technologies. The thematic focus of the next innovation report will thus be placed on the status of relevant vehicle engineering and an intensive appraisal of the further development expected in the coming years. On this basis, scientifically founded recommendations can then be offered to show how the increased accident risk associated with technical progress can be alleviated through a demanding system of novice driver preparation.

# Bibliography


Allen, R.W., Park, G.D., Cok, M.L. & Fiorentino, D. (2007). *The effect of driving simulator fidelity on training effectiveness.* Iowa City: Driving Simulator Conference.

Amelang. M. & Zielinski, W. (2002). *Psychologische Diagnostik und Intervention.* Heidelberg: Springer.

Arendasy, M.E., Hergovich, A. & Sommer, M. (2005). Dimensionality and construct validity of a video-based, objective personality test for the assessment of willingness to take risks in road traffic. *Psychological Reports, 97(1)*. 309-20.

Bailly, B., Bellet, T. & Goupil, C. (2003). *Drivers' mental representations: experimental study and training perspectives*. Paper presented at the First International Conference on Driving Behaviour and Training, Stratford-upon-Avon, UK.

Barthelmess, W. (1976). Erfahrung und Risiko. Ein theoretischer Beitrag zur Erfahrungsbildung des Verkehrsteilnehmers. *Sicherheit + Zuverlässigkeit in Wirtschaft, Betrieb, Verkehr, TÜ, Bd. 7/8*, 255-298.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

BASt-Expertengruppe (2012, unpublished manuscript). *"Fahranfängervorbereitung": Rahmenkonzept zur Weiterentwicklung der Fahranfängervorbereitung in Deutschland.* Bundesanstalt für Straßenwesen, Referat U4 "Fahrausbildung, Kraftfahrerrehabilitation", Bergisch Gladbach.

Benda, H. v. (1985). Die Häufigkeit von Verkehrssituationen. Bericht zum Forschungsprojekt 7320/2 der Bundesanstalt für Straßenwesen, Bereich Unfallforschung. Bergisch Gladbach: BASt.

Benda, H. v. & Hoyos, C.G. (1983). Estimating hazards in traffic situations. *Accident Analysis & Prevention, 15 (1)*, 1-9.

Bernotat, R. & Käppler, W.-D. (1985). Wirksystem Fahrer-Fahrzeug-Umwelt. In K. Rompe (Hrsg.), *Verkehrssicherheit und Wirksystem Fahrer-Fahrzeug-Umwelt*. Kolloquium d. Inst. für Verkehrssicherheit im TÜV Rheinland e.V., Direktionsbereich Kraftfahrzeugverkehr; Veranst. d. TÜV-Akad. Rheinland am 29. November 1984 in Köln-Poll (pp. 13-44). Cologne: Verlag TÜV Rheinland.

Biedinger, J., Friedel, T., Genschow, J., Möhlke, S., Rüdel, M., Weiße, B. & Wagner, W. (2012). *Tätigkeitsbericht zum Betrieb, zur Evaluation und zur Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2011*. Dresden: TÜV | DEKRA arge tp21.

Böhne, A., Genschow, J., Rüdel, M., Schubert, T., Steuer, M., Teichert, C., Wagner, W. & Weiße, B. (2014). *Tätigkeitsbericht zum Betrieb, zur Evaluation und zur Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2013*. Dresden: TÜV | DEKRA arge tp 21.

Böhne, A., Genschow, J., Rüdel, M., Schubert, T., Steuer, M., Wagner, W. & Weiße, B. (2013). *Tätigkeitsbericht zum Betrieb, zur Evaluation und zur Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2012*. Dresden: TÜV | DEKRA arge tp21.

Bönninger, J. (2014a). *Pluralisierung Fahrzeugkonzepte – Mobilitätsrechte 21*. Unpublished presentation. Radeberg.

Bönninger, J. (2014b). Wem gehören die Daten im Fahrzeug? Das moderne Fahrzeug – Messgerät, Steuergerät, Datenspeicher. *Zeitschrift für Schadensrecht (4)*, 184–189.

Bönninger, J., Kammler, K. & Sturzbecher, D. (Hrsg.). (2009). *Die Geschichte der Fahrerlaubnisprüfung in Deutschland*. Bonn: Kirschbaum.

Bönninger, J. & Sturzbecher, D. (2005). *Optimierung der Fahrerlaubnisprüfung. Ein Reformvorschlag für die theoretische Fahrerlaubnisprüfung.* Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 168. Bremerhaven: Wirtschaftsverlag NW.

Borowsky, A., Oron-Gilad, T. & Parmet, Y. (2009). Age and skill differences in classifying hazardous traffic scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 12 (4)*, 277–287.

Borowsky, A., Shinar, D. & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis & Prevention, 42 (4),* 1240–1249.

Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.

Brauckmann, J., Hähnel, R., & Mylius, G. (2006). *Der Kraftfahrsachverständige. Die Technischen Prüfstellen und ihre amtlich anerkannten Sachverständigen und Prüfer für den Kraftfahrzeugverkehr.* Bonn: Kirschbaum Fachverlag für Verkehr und Technik.

Bredow, B. (2014). *Die Zukunft der Gefahrenlehre in der Fahrschulausbildung – Evaluation des pädagogisch-psychologischen Verkehrssicherheitsprojekts "Regio-Protect 21".* Hannover: Degener.

Bredow, B. & Sturzbecher, D. (in press). *Ansätze zur Optimierung der Fahrschulausbildung*. Abschlussbericht zum Forschungsprojekt FE 82.515/2011 im Auftrag der Bundesanstalt für Straßenwesen.

Brown, I.D. (1982). Exposure and experience are a confounded nuisance in research on driver behaviour. Accident Analysis & Prevention, 14 (5), 345-352.

Brown, I.D. & Groeger, J.A. (1988). Risk perception and decision taking during the transition between novice and experienced driver status. *Ergonomics, 31 (4),* 585-597.

Bundesministerium für Verkehr und digitale Infrastruktur (2014-12-16). *Runder Tisch "Automatisiertes Fahren"*. 2. Sitzung des Plenums. Berlin.

Carpentier, A., Wang, W., Jongen, E.M.M., Hermans, E. & Brijs, T. (2012). *Training hazard perception of young novice drivers. A driving simulator study*.

Castro, C. (Ed.). (2009). *Human factors of visual and cognitive performance in driving*. Boca Raton, Fla.: CRC Press.

Catchpole, J., Congdon, P., & Leadbeatter, C. (2001). *Implementation of Victoria's new hazard perception test.* Road Safety Research: Policing And Education Conference. Melbourne, Victoria, Australia.

Catchpole, J . & Leadbeatter, C. (2000). *Redevelopment of Victoria's hazard perception test.* Road Safety Research, Policing and Education Conference, 2000. Brisbane, Queensland, Australia.

Chapman, P., Crundall, D. E., Phelps, N. & Underwood, G. (2002). Police drivers' visual search in hazardous situations. In: Department for Transport (Ed.), Behavioural Research in Road Safety. Twelfth Seminar (pp. 57-67). London.

Chapman, P., Underwood, G. & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 5(2),* 157-167.

Chapman, P.R. & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception, 27*, 951 -964.

Clarke, D.D., Ward, P., Bartle, C. & Truman, W. (2004). *In-depth study of motorcycle accidents. Road safety research report no. 54.* Department for Transport: London.

Clarke, D.D., Ward, P., Bartle, C. & Truman, W. (2006). Young driver accidents in the UK: The influence of age, experience, and time of day. *Accident Analysis & Prevention, 38 (5),* 871-878.

Clarke, D.D., Ward, P. & Truman, W. (2005). Voluntary risk taking and skill deficits in young driver accidents in the UK. *Accident Analysis & Prevention, 37 (3),* 523-529.

Collins, A.M. & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychological Review (83)*, 407-428.

Congdon, P. & Cavallo, A. (1999). *Validation of the Victorian hazard perception test*. Road Safety Research: Policing And Education Conference. Canberra, Australia.

Continental (2013). *Continental Mobilitätsstudie 2013*. Accessed on 24.06.2015. URL: https://www.infas.de/fileadmin/user_upload/Praxisforum/infas_InnoZ_Automat_Fahren_Continental.pdf.

Crick, N.R. & Dodge, K.A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74-101.

Crundall, D., Andrews, B., van Loon, E. & Chapman, P. (2010). *Commentary training improves responsiveness to hazards in a driving simulator.* Nottingham: Accident Research Unit.

Crundall, D., Chapman, P., Trawley, S., Collins, L., van Loon, E., Andrews, B. & Underwood, G. (2012). Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard. *Accident Analysis & Prevention, 45*, 600-609.

Crundall, D. & Underwood, G. (2011). Visual Attention While Driving. Measures of Eye Movements Used in Driving Research. In B.E. Porter (Ed.), *Handbook of Traffic Psychology*. Elsevier.

Crundall, D., Underwood, G. & Chapman, P. (1999). Driving experience and the functional field of view. *Perception, 28 (9),* 1075-1087.

Crundall, D., Underwood, G. & Chapman, P. (2002). Attending to the peripheral world while driving. *Applied Cognitive Psychology, 16 (4),* 459-475.

Cuvenhaus, H., Genschow, J. & Sturzbecher, D. (2013). *Wissenschaftliches Gutachten – Optimierung der Bewertungssystematik für die Theoretische Fahrerlaubnisprüfung*. Kremmen: IPV.

de Winter, J.C.F., Wieringa, P.A., Dankelmann, J., Mulder, M., van Paasen, M.M. & de Groot, S. (undated). *Driving simulator fidelity and training effectiveness.* Delft: University.

DeCraen, S. (2010). The X-factor. A longitudinal study of calibration in young novice drivers. SWOV, Leidschendam

Deery, H. A. (1999). Hazard and Risk Perception among Young Novice Drivers. *Journal of Safety Research, 30 (4),* 225-236.

Donges, E. (2009). Fahrerverhaltensmodelle. In H. Winner, S. Hakuli & G. Wolf (Hrsg.), *Handbuch Fahrerassistenzsysteme. Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (pp. 15-23). Wiesbaden: Vieweg & Teubner.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Facotrs 37 (1),* 32-64.

Engelkamp, J. (2004). Gedächtnis für Bilder. In K. Sachs-Hombach & K. Rehkämper (Hrsg.), *Bild – Bildwahrnehmung – Bildverarbeitung*. Deutscher Universitätsverlag.

Ewert, U. & Steiner, K. (2013). *Fahrsimulatoren für Aus- und Weiterbildung. Beratungsstelle für Unfallverhütung (bfu) – Faktenblatt Nr. 11.* Berne: bfu.

Fack, D. (2000). *Automobil, Verkehr und Erziehung. Motorisierung und Sozialisation zwischen Beschleunigung und Anpassung 1885-1945.* Opladen: Leske & Budrich.

Fahrschüler-Ausbildungsordnung (2012), geändert durch Artikel 3 der Verordnung vom 16. April 2014 (BGBl. I p. 348)

Fahrschüler-Ausbildungsordnung (FahrschAusbO), Date of issue: 19.06.2012 (Accessed on 23.06.2015, URL: http://www.gesetze-im-internet.de/fahrschausbo_ 2012/).

Falkmer, T. & Gregersen, N.P. (2001). Fixation patterns of learner drivers with and without cerebral palsy (CP) when driving in real traffic environments. *Transportation Research Part F: Traffic Psychology and Behaviour, 4 (3),* 171-185.

Fastenmeier, W. (Hrsg.). (1995). *Autofahrer und Verkehrssituation – Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Straßenverkehrssysteme* (Mensch-Fahrzeug-Umwelt, Bd. 33). Cologne: Verlag TÜV Rheinland.

Fisher, D.L. (2008). *Evaluation of PC-based novice driver risk awareness.* Amherst: NHTSA.

Fisher, D.L., Pollatsek, A.P. & Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention, 12 (Suppl 1),* 25-29.

Fisseni, H.J. (1990). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe

Fournier, H. (1901). Wie man fahren soll. *Allgemeine Automobil-Zeitung, 47(2),* 6-7.

Friedel, T., Mörl, S. & Rüdel, M. (2012). *Bericht zur Machbarkeitsstudie für die Entwicklung und Ersterprobung eines elektronischen Prüfprotokolls zur Optimierung der Praktischen Fahrerlaubnisprüfung*. Dresden: TÜV | DEKRA arge tp 21.

Friedel, T. & Rüdel, M. (2010). *Erprobung innovativer Aufgaben im Rahmen der regulären Theoretischen Fahrerlaubnisprüfung – Methodisches Vorgehen*. Dresden: TÜV | DEKRA arge tp 21.

Friedel, T. & Rüdel, M. (2011). *Wiedererkennung variierter statischer Bilder von Verkehrssituationen* (Vorläufiger Ergebnisbericht). Umpublished manuscript. Dresden: TÜV | DEKRA arge tp 21.

Friedel, T., Weiße, B., Genschow, J. & Schmidt, A. (2011). Entwicklung innovativer Aufgabenformate mit computergenerierten Szenarien. In TÜV | DEKRA arge tp21 (Hrsg.), *Innovationsbericht zur Optimierung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2009/2010*. Dresden: TÜV | DEKRA arge tp21

Frisch, S., Burkert, M. & Genschow, J. (2013). *Untersuchung zur Äquivalenz von herkömmlichen und PC-generierten Abbildungen in Prüfungsaufgaben* (Berichte zur Evaluation der Theoretischen Fahrerlaubnisprüfung). Oberkrämer: IPV.

Frisch, S., Teichert, C. & Genschow, J. (2013). *Untersuchung zur Äquivalenz von Prüfungsaufgaben mit variierten Abbildungen* (Berichte zur Evaluation der Theoretischen Fahrerlaubnisprüfung). Oberkrämer: IPV.

Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.

Genschow, J., Krampe, A. & Weiße, B. (2011). Evaluation und Weiterentwicklung des Prüfungsverfahrens. In TÜV I DEKRA arge tp 21 (Hrsg.), *Innovationsbericht zur Optimierung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2009/2010*. Dresden: TÜV | DEKRA arge tp 21

Genschow, J., Sturzbecher, D. & Willmes-Lenz, G. (2013). *Fahranfängervorbereitung im internationalen Vergleich.* Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 234. Bremerhaven: Wirtschaftsverlag NW.

Gerster, B. (2013). *Fahrer Assistenz Systeme – Wohin führt der Weg*. AGU Seminar 2013. Zürich.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser & M. J. Farr (Eds.), The nature of expertise (pp. xv-xxviii). Hillsdale: NJ: Erlbaum.

Glowalla, P. & Jordan, C. (2014). *Diskussionsgrundlage zum Erwerb einer Mofa-Prüfbescheinigung und einer Fahrerlaubnis der Klasse AM.* Unpublished manuscript. Berlin.

Goldstein, B.E. (1997). *Wahrnehmungspsychologie.* Heidelberg: Spektrum.

Grattenthaler, H., Krüger, H.-P. & Schoch, S. (2009). *Bedeutung der Fahrpraxis für den Kompetenzerwerb beim Fahrenlernen.* Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 201. Bremerhaven: Wirtschaftsverlag NW.

Grayson, G.B., Maycock, G., Groeger, J.A., Hammond, S.M. & Field, D.T. (2003). *Risk, hazard perception and perceived control* (TRL Report 560). Crowthorne, Berkshire: TRL Limited.

Grayson, G.B. & Sexton, B.F. (2002). *The development of hazard perception testing.* TRL Report 558. Crowthorne, Bershire: TRL Limited.

Gruber, H. & Mandl, H. (1996). Das Entstehen von Expertise. In J. Hoffmann & W. Kintsch (Hrsg.), *Enzyklopädie der Psychologie, C/II/7* (pp. 583-615). Göttingen: Hogrefe.

Gürten, J., Neumeier R. & Wiegand W. (1987). *Punktefrei und sicher fahren: Ein Trainingsprogramm mit Nachbetreuungsphase zur Verbesserung der Selbstkontrolle beim Fahren.* Cologne: Verlag TÜV Rheinland.

Hagen, K., Schulze, C. & Schlag, B. (2012). *Verkehrssicherheit von schwächeren Verkehrsteilnehmern im Zusammenhang mit dem geringen Geräuschniveau von Fahrzeugen mit alternativen Antrieben.* Forschungsvereinigung Automobiltechnik e.V., FAT-Schriftenreihe 245.

Hampel, B. (1977a). *Erprobung eines audiovisuellen Prüfungssystems des TÜV Rheinland − Ergebnisse und Konsequenzen*. Schriftenreihe des Medizinisch-Psychologischen Instituts des TÜV Rheinland: Mensch – Fahrzeug – Umwelt, Bd. 4, Entwicklung und Konzepte für die Fahrerlaubnisprüfung (pp. 57-91). Cologne: TÜV Rheinland.

Hampel, B. (1977b). *Möglichkeiten zur Standardisierung der praktischen Fahrerlaubnisprüfung*. Bericht zum Forschungsauftrag 7516 der Bundesanstalt für Straßenwesen. Cologne: Technischer Überwachungs-Verein Rheinland e.V.

Hampel, B. (1979). *Bericht über Ergebnisse der Demonstration eines audio-visuellen Verkehrskenntnistests anlässlich der Internationalen Verkehrsausstellung 1979 in Hamburg*. Cologne: TÜV Rheinland.

Harrison, G.W., Triggs, T.J. & Pronk, Nicola J. (1999). *Speed and young drivers: developing countermeasures to target excessive speed behaviours amongst young drivers* (Report No. 159). Victoria.

Hartig, J. & Klieme, E. (2007). *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (Vol. 20). Bundesministerium für Bildung und Forschung (BMBF).

Helman, S. (2008). Situational judgement in driver training and assessment: A literature review. In Transport Research Laboratory (TRL). (Ed.), *Published Project Report PPR312*. Accessed on 26.10.2014. URL: http://www.rsa.ie/documents/learner%20drivers/gdl/situational%20judgement%20in%20driver%20training%20and%20assessment.pdf.

Hergovich, A., Arendasy, M.E., Sommer, M. & Bognar, B. (2007). The Vienna Risk-Taking Test – Traffic. A New Measure of Road Traffic Risk-Taking. *Journal of Individual Differences 2007; Vol. 28(4)*,198-204.

Hilz, J., Malone, S. & Brünken, R. (2014). *Development and evaluation of a computer based training program for driver education.* Poster presented at the 2nd International Conference on Human Factors in Transportation 2014, Krakow, Poland.

Hodges, N.J., Huys, R., & Starkes, J.L. (2007). Methodological review and evaluation of research in expert performance in sport. In G. Tenenbaum & R.C. Eklund (Eds.), *Handbook of Sport Psychology* (Vol. 3, pp. 161-183). Hoboken, NJ: John Wiley & Sons.

Höffler, T.N. & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*, 722-738.

Hoffmann, S. & Buld, S. (2006). Darstellung und Evaluation eines Trainings zum Fahren in der Fahrsimulation. In VDI-Gesellschaft Fahrzeug- und Verkehrstechnik |f ED (Hrsg.), *Integrierte Sicherheit und Fahrerassistenzsysteme* (pp. 113-132). Düsseldorf: VDI Verlag.

Hole, G.J. (2007). *The Psychology of Driving*. New Jersey: Erlbaum.

Horswill, M.S., Anstey, K.J., Hatherly, C.G. & Wood, J.M. (2010). The crash involvement of older drivers is associated with their hazard perception latencies. *Journal of the International Neuropsychological Society, 16 (05)*, 939-944.

Horswill, M.S. & McKenna, F.P (2004). Drivers hazard perception ability: situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness. Theory and application* (pp. 155-175). Ashgate Publishing.

Huestegge, L., Skottke, E.-M., Anders, S., Müsseler, J. & Debus, G. (2010). The development of hazard perception: Dissociation of visual orientation and hazard processing. *Transportation Research Part F: Traffic Psychology and Behaviour, 13 (1),* 1-8.

IPCS (2004). *Risk assessment terminology – Part 1 and Part 2*. Geneva, Switzerland: World Health Organization.

Irish Drivers Education Association (2006). *Steer Clear-Curriculum. Driver Education – Learning for Life.* Steer Wicklow: IDEA Ltd.

Isler, R.B., Starkey, N.J., & Williamson, A.R. (2009). Video-based road commentary training improves hazard perception of young drivers in a dual task. *Accident Analysis & Prevention, 41(3),* 445-452.

Itti, L. & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research 40*, 1489-1506.

Jurecka, A. & Hartig, J. (2007). Computer- und netzwerkbasiertes Assessment. In Bundesministerium für Bildung und Forschung (BMBF). (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (pp. 37-48). Bonn, Berlin: Bundesministerium für Bildung und Forschung.

Kahnemann, D. (1970). *Remarks on attention control*. Acta psychologica, 33, Attention and performance, 3rd edition.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise* (unveränderte Auflage). Bonn, Berlin: Bundesministerium für Bildung und Forschung. Accessed on 16.11.2012. URL: http://www.bmbf.de/pub/zur_ entwicklung_nationaler_bildungsstandards.pdf.

Klieme, E. & Leutner, D. (2006). *Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen*. Überarbeitete Fassung des Antrags an die DFG auf Einrichtung eines Schwerpunktprogramms. Accessed on 23.08.2011. URL: http://kompetenzmodelle.dipf.de/pdf/rahmenantrag.

Kluwe, R. & Spada, H. (1981). Wissen und seine Veränderung: Einige psychologische Beschreibungsansätze. In K. Foppa & R. Groner (Hrsg.), *Kognitive Strukturen und ihre Entwicklung*. Stuttgart: Huber.

Koch, I. (2008). Mechanismen der Interferenz in Doppelaufgaben. *Psychologische Rundschau, 59 (1)*, 24-32.

Koeppen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current issues in competence modeling and assessment. *Journal of Psychology, 216 (2)*, 61-73.

König, W. (2012). Nutzergerechte Entwicklung der Mensch-Maschine- Interaktion von Fahrerassistenzsystemen. In H. Winner, S. Hakuli & G. Wolf (Hrsg.), *Handbuch Fahrerassistenzsysteme. Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (2., korrigierte Auflage). ATZ/MTZ-Fachbuch. Wiesbaden: Vieweg + Teubner.

Kötter, S. & Nordmann, E. (1987). Diagnostische Beobachtungsverfahren. In M. Cierpka (Hrsg.), *Familiendiagnostik*. Berlin: Springer.

Krüger, H.P. (2010). *"Ursachenanalyse"*. AK VII Unfallrisiko "Junge Fahrer". Vortrag zum 48. Deutschen Verkehrsgerichtstag in Goslar, 27.-29.01.2010 .

Kühn, M. (2008). *Analyse des Motorradunfallgeschehens. Unfallforschung kompakt*. Berlin: GDV. Unfallforschung der Versicherer.

Kuiken, M. & Twisk, D. (2001). *Safe driving and the training of calibration*. SWOV, Leidschendam.

Laapotti, S. & Keskinen, E. (1998). Differences in fatal loss-of-control accidents between young male and female drivers. *Accident Analysis & Prevention, 30 (4),* 435-442.

Lazarus, R.S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Lesch, M. F. (2008). Warning symbols as reminders of hazards: Impact of training. *Accident Analysis & Prevention, 40 (3)*, 1005-1012.

Leutner, D., Brünken, R. & Willmes-Lenz, G. (2009). Fahren lernen und Fahrausbildung. In H.P. Krüger (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete, Serie VI Verkehrspsychologie, Bd. 2 Anwendungsfelder der Verkehrspsychologie* (pp. 1-79). Göttingen: Hogrefe.

Lienert, G.A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.

Lonero, L., Clinton, K., Brock, J., Wilde, G., Laurie, I. & Black, D. (1995). *Novice Driver Education Curriculum Outline.* Foundation for Traffic Safety Research. Accessed on 26.10.2010. URL: http://www.aaafts.org/resources/index.cfm?button=lonaro.

Ludwig Bölkow Systemtechnik GmbH (2006). MAN Brennstoffzellenbus. Accessed on 05.08.2014. URL: http://www.brennstoffzellenbus.de/bus/.

Luhmann, N. (1981). Schematismen der Interaktion. In N. Luhmann (Hrsg.), *Soziologische Aufklärung 3. Soziales System, Gesellschaft, Organisation* (pp. 81–100). Wiesbaden: VS Verlag für Sozialwissenschaften.

Luhmann, N. (2012). *Soziale Systeme. Grundriß einer allgemeinen Theorie* (15. Aufl). Frankfurt am Main: Suhrkamp.

Luniak, P., Mörl, S. & Friedel, T. (2014). *Zwischenbericht zum Stand der Umsetzung des BASt-Projekts "Revision zu einer optimierten Praktischen Fahrerlaubnisprüfung"* (FE 82.0529/2011). Kremmen: Unpublished manuscript.

Lyon, J., Borkenhagen, D., Scialfa, C., Deschêne, M. & Horswill, M. (2011). *Developing a North American static hazard perception test.* Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, University of Iowa.

Malone, S. (2012). *Computerbasierte Messung von Teilaspekten der Fahrkompetenz. Besonderheiten des Expertiseerwerbs beim Autofahren.* Saarbrücken: Universität des Saarlandes.

Malone, S., Biermann, A., Brünken, R. & Buch, S. (2012). *Neue Aufgabenformate in der Fahrerlaubnisprüfung.* Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 222, Bremerhaven: Wirtschaftsverlag NW.

Malone, S., & Brünken, R. (2013a). *Validierung handlungsnaher Antwortformate für dynamische Aufgabenformate in der Fahrerlaubnisprüfung*. Unpublished. Dresden: TÜV | DEKRA arge tp 21.

Malone, S., & Brünken, R. (2013b). Assessment of driving expertise using multiple choice questions including static vs. animated presentation of driving scenarios. *Accident Analysis & Prevention, 51*, 112-119.

Marek, J. & Sten, T. (1972). *Driver behavior, training, and traffic environment.* Trondheim: University.

Mason, M. (2003). *Validation of the hazard perception test for new drivers*. Bedford: Cranfield University.

Maycock, G., Lockwood, C.R. & Lester, J.F. (1991). *The accident liability of car drivers*. Research Report 315. Crowthorne, Berkshire : TRL.

McGwin, J. & Brown, D.B. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis & Prevention, 31 (3)*, 181-198.

McKenna, F.P., & Crick, J.L. (1994a). *Developments in hazard perception*. Final Report: Department of Transport (UK).

McKenna, F.P & Crick, J.L. (1994b). *Hazard Perception in Drivers: A Methodology for Testing and Training.* Contractor Report 313, Crowthorne, Berkshire : TRL.

McKenna, F.P. & Crick, J. (1997). *Developments in hazard perception* (TRL report Nr. 297). Crowthorne (UK): Transport Research Laboratory.

McKenna, F.P. & Horswill, M.S. (1997). Differing conceptions of hazard perception. In G.B. Grayson (Ed.), *Behavioural Research in Road Safety VII* (pp. 74-81). Crowthorne: TRL.

McKnight, A.J. & Adams, B.B. (1970a). *Driver Education Task Analysis. Volume I: Task Desriptions.* Alexandria, VA: Human Resources Research Organization.

McKnight, A.J. & Adams, B.B. (1970b). *Driver Education Task Analysis. Volume II: Task Analysis Methods. Final Report*. Alexandria, VA: Human Resources Research Organization.

McKnight, A.J. & Adams, B.B. (1971). *The Development of Driver Education Objectives Through an Analysis of the Driving Task*. Alexandria, VA: Human Resources Research Organization.

McKnight, A.J. & Hundt, A.G. (1971). Driver Education Task Analysis. Volume III: Instructional Objectives. Final Report. Alexandria, VA: Human Resources Research Organization.

McKnight, A.J. & McKnight, A.S. (2003). Young novice drivers: careless or clueless? *Accident Analysis & Prevention*, 921-925.

Mead, M. (1971). *Der Konflikt der Generationen. Jugend ohne Vorbild?* Olten: Walter-Verlag.

Meir, A., Borowsky, A., Oron-Gilad, T., Parmat, Y. & Shinar, D. (2010). *Towards developing a hazard perception training program for enhancing young-inexperienced drivers' abilities.* Be'er Sheva: Ben Gurion University.

Mills, K.L., Hall, R.D., McDonald, M., & Rolls, G.W.P. (1998). *The effects of hazard perception training on the development of novice driver skills.* Report to Department Environment, Transport & Regions.

Moosbrugger, H. & Kelava, A. (2008). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer

Mörl, S., Kleutges, Ch. & Rompe, K. (2009). Die Entwicklung der Fahrerlaubnisprüfung in der Bundesrepublik bis 1989. In D. Sturzbecher, J. Bönninger & K. Kammler (Hrsg.), *Die Geschichte der Fahrerlaubnisprüfung in Deutschland.* Dresden: TÜV | DEKRA arge tp 21.

Möser, K. (2004). "Der Kampf des Automobilisten mit seiner Maschine" – Eine Skizze der Vermittlung der Autotechnik und des Fahrenlernens im 20. Jahrhundert. In L. Bluma, K. Pichol & W. Weber (Hrsg.). *Technikvermittlung und Technikpopularisierung. Historische und didaktische Perspektiven.* Münster.

Mourant, R.R. & Rockwell, T.H. (1972). Strategies of Visual Search by Novice and Experienced Drivers. *Human Factors, 14 (4)*, 325-335.

Munsch, G. (1973). *Dynomen-Lehre. Eine psychologisch-pädagogische Studie über die Notwendigkeit und die Möglichkeiten des Trainings der "Vorahnung" kritischer Verkehrslagen*. Technischer Überwachungs-Verein Bayern e.V.

Nakamura, C. & Zeng-Treitler, Q. (2012). A taxonomy of representation strategies in iconic communication. *International Journal of Human-Computer Studies, 70 (8)*, 535-551.

Neale, V.L., Dingus, T.A., Klauer, G.S., Sudweeks, J. & Goodman, M. (2005). *An overview of the 100-car naturalistic study and findings* (DOT HS Publication 05- 0400).

Nelson, K. (Ed.). (1986). *Event knowledge. Structure and function in development* (pp. 97-118). Hillsdale, NJ: Erlbaum.

Nett, N. & Frings, C. (2015). Signalentdeckungstheorie. In M.A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie*. Accessed on 03.06.2014. URL: https://portal.hogrefe.com/dorsch/signal-entdeckungstheorie/.

Noordzij, P.C., Forke, E., Brendicke, R. & Chinn, B.P. (2001). *Integration of needs of moped and motorcycle riders into safety measures*. Review and statistical analysis in the framework of the European research project PROMISING, Work Package 3. Accessed on 23.04.2014. URL: http://ec.europa.eu/transport/roadsafety_library/publications/promising_deliverable_3.pdf

O'Regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24 (05)*, 939-1031.

Ostwald, K. (1931). Jung-Deutschland und die Kraftfahrt. *Allgemeine Automobil-Zeitung 32(8),* 40.

Parsons, T. & Shils, E.A. (Eds.) (1951). *Toward a General Theory of Action*. Cambridge, MA: Harvard University Press.

Pelz, D.C. & Krupat, E. (1974). Caution Profile and Driving record of Undergraduate Males. *Accident Analysis and Prevention, 6*, 45-58.

Petzold, T., Weiß, T., Franke, T., Krems, J.F. & Bannert, M. (2011). *Unterstützung der Fahrausbildung durch Lernsoftware*. Berichte der Bundesanstalt für Straßenwesen. Heft M 219. Bremerhaven: Wirtschaftsverlag NW.

Pew, R. W. (2008). The state of situation awareness measurement: Heading toward the next century. In M.R. Endsley & D.J. Garland (Hrsg.), *Situation awareness analysis and measurement* (Reprinted) (pp. 33-47). Boca Raton: CRC Press.

Pollatsek, A., Fisher, D.L. & Pradhan, A.K. (2006). Identifying and Remediating Failures of Selective Attention in Younger Drivers. *Current Directions in Psychological Science, 15,* 255-259.

Prücher, F. (2006). Computer based training in driver education and current developments on a computer assisted driving test in Germany. In *Proceedings of the Conference on European guidelines for the application of new technologies for driver training and education,* Madrid, April 25-26, Annex 3. HUMANIST, Universidad Politécnica de Madrid.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13(3),* 257-266.

Rath, V. & Bozem, K. (2013). Technologietrends Automotive und deren energiewirtschaftliche Implikationen. In K. Bozem, A. Nagl & C. Rennhak (Hrsg.), *Energie für nachhaltige Mobilität. Trends und Konzepte* (pp. 73-115). Wiesbaden: Springer Fachmedien Wiesbaden.

Regan, M.A., Triggs, T.J. & Godey, S.T. (2000). *Simulator-based evaluation of the DriveSmart novice driver CD-ROM training product.* Monash: MUARC.

Reichart, G. (2001). *Menschliche Zuverlässigkeit beim Führen von Kraftfahrzeugen*. Dissertation (Als Ms. gedr). Düsseldorf: VDI-Verl.

Rensink, R.A. (2002). Change Detection. *Annual Review of Psychology, 53*, 245-277.

Richtlinie für die Ausbildung von Fahrschülern für den Kraftfahrzeugverkehr (1971).

Rost, J. (2004). *Lehrbuch Testtheorie-Testkonstruktion*. Berne: Huber.

Rüdel, M., Sturzbecher, D., Genschow. J. & Weiße, B. (2011). Entwicklung innovativer Aufgabentypen und Prüfungsformen zur erweiterten Fahrkompetenzerfassung. In TÜV | DEKRA arge tp 21 (Hrsg.), *Innovationsbericht zur Optimierung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2009/2010*. Dresden: TÜV | DEKRA arge tp 21.

Russell, E. (2003). *National Handbook of Traffic Control Practices for Low Volume Rural Roads and Small Cities. VOLUME I: Low-Volume Roads.* Mack Blackwell Transportation Center. Kansas State University, First National Edition.

Sagberg, F. & Bjørnskau, T. (2006). Hazard perception and driving experience among novice drivers. *Accident Analysis and Prevention, 38,* 407-414.

Schade, F.-D. (2001). *Daten zur Verkehrsbewährung von Fahranfängern*. Reanalyse von Rohdaten der Untersuchung Hansjosten, E. und Schade, F.-D. (1997. Legalbewährung von Fahranfängern. Berichte der BASt, Reihe Mensch und Sicherheit, Heft M71). Unpublished manuscript, Flensburg: Kraftfahrt-Bundesamt.

Schank, R.C. (1975). Concepts for representing mundane reality in plans. In D.G. Bobrow & A. Collins (Eds.), *Representation and understanding*. New York: Academic Press.

Schank, R.C. & Abelson, R.P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

Schlag, B. (2009). Visuelle Wahrnehmung und Informationsaufnahme im Straßenverkehr. In Schlag, B., Petermann, I., Weller, G. & Schulze, C. (Hrsg.), *Mehr Licht – mehr Sicht – mehr Sicherheit? Zur Wirkung verbesserter Licht- und Sichtbedingungen auf das Fahrerverhalten*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Schlag, B., Ellinghaus, D. & Steinbrecher, J. (1986). *Risikobereitschaft junger Fahrer*. Bremerhaven: Wirtschaftsverl.

Schlag, B., Petermann, I., Weller, G. & Schulze, C. (2009). Mehr Licht – mehr Sicht – mehr Sicherheit? Wiesbaden: VS Verlag für Sozialwissenschaften.

Schlick, T. (2013). *Elektromobilität im internationalen Vergleich*. Frankfurt am Main.

Schneider, W. (o. V.). (1976). Fahrschule: Lern mal was fürs Überleben. *Der Spiegel, 18,* 70-81.

Schubert, W. (2000). Das Gesetz über den Verkehr mit Kraftfahrzeugen vom 3.5.1909. *Zeitschrift der Savigny-Stiftung für Rechtsgeschichte, Germanistische Abteilung, 117,* 238-289.

Schulz-Zander, R. (2005). Veränderung der Lernkultur mit digitalen Medien im Unterricht. In H. Kleber (Hrsg.), *Perspektiven der Medienpädagogik in Wissenschaft und Bildungspraxis* (pp. 125-140). Munich: Kopaed-Verlag.

Scialfa, C.T., Borkenhagen, D., Lyon, J. & Deschênes, M. (2013). A comparison of static and dynamic hazard perception tests. *Accident Analysis & Prevention, 51*, 268-273.

Scialfa, C.T., Borkenhagen, D., Lyon, J., Deschênes, M., Horswill, M.S. & Wetton, M. (2012). The effects of driving experience on responses to a static hazard perception test. *Accident Analysis and Prevention, 45*, 547-553.

Scialfa, C.T., Deschênes, M.C., Ference, J., Boone, J., Horswill, M.S. & Wetton, M. (2011). A hazard perception test for novice drivers. *Accident Analysis & Prevention, 43 (1),* 204-208.

Seidl, J. (1990). *Die Kommentarmethode in der Fahrschulausbildung. Psychologische Untersuchungen zum Einsatz des Verbalisierens in der Ausbildung von Kraftfahrzeugführern.* Dresden: Technische Universität Dresden.

Seidl, J. & Hacker, W. (1991). Verbalisierung in der Fahrschulausbildung. Psychologische Untersuchungen zum Einsatz der Kommentarmethode bei der fahrpraktischen Ausbildung von Kraftfahrzeugführern. *Zeitschrift für Verkehrssicherheit 37 (3),* 109-116.

Silbereisen, R.K. (1987). Soziale Kognition – Entwicklung von sozialem Wissen und Verstehen. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (pp. 696-736). Munich: Psychologie Verlags Union.

Sivak, M. (1996). The information that drivers use: is it indeed 90% visual? *Perception, 25 (9),* 1081-1089.

Smith, S.S., Horswill, M.S., Chambers, B. & Wetton, M. (2009). Hazard perception in novice and experienced drivers: The effects of sleepiness. *Accident Analysis & Prevention, 41 (4)*, 729-733.

Soliday, S.M. (1974). Relationship between age and hazard perception in automobile drivers. *Perceptual and Motor Skills (39),* 335-338.

Spelke, E., Hirst, W. & Neisser, U. (1976). Skills of divided attention. *Cognition, 4*, 215-230.

Standing, L. (1973). Learning 10000 pictures. In: Quarterly Journal of Experimental Psychology, 25:2, 207-222

Statistisches Bundesamt (Hrsg.). (2013). *Verkehrsunfälle. Unfälle von 18- bis 24-Jährigen im Straßenverkehr.* Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2014). *Verkehr 2013. Verkehrsunfälle.* Fachserie 8, Reihe 7. Wiesbaden: Statistisches Bundesamt.

Straßenverkehrsgesetz (StVG). Accessed on 03.06.2015. URL: http://www.gesetze-im-internet.de/bundesrecht/stvg/gesamt.pdf.

Straßenverkehrs-Ordnung (StVO), Date of issue: 06.03.2013. Accessed on 23.06.2015. URL: http://www.gesetze-im-internet.de/stvo_2013/.

Sturzbecher, D. (2010). Methodische Grundlagen der praktischen Fahrerlaubnisprüfung. In D. Sturzbecher, J. Bönninger & M. Rüdel (Hrsg.), *Praktische Fahrerlaubnisprüfung – Grundlagen und Optimierungsmöglichkeiten.* Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 215 (pp. 17-38). Bremerhaven: Wirtschaftsverlag NW.

Sturzbecher, D., Bönninger, J. & Rüdel, M. (Hrsg.). (2008). *Optimierung der praktischen Fahrerlaubnisprüfung. Methodische Grundlagen und Möglichkeiten der Weiterentwicklung*. Dresden: TÜV | Dekra arge tp 21.

Sturzbecher, D., Bönninger, J. & Rüdel, M. (Hrsg.). (2010). *Praktische Fahrerlaubnisprüfung – Grundlagen und Optimierungsmöglichkeiten*. Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 215. Bremerhaven: Wirtschaftsverlag NW.

Sturzbecher, D., Kammler, K. & Bönninger, J. (2005). Möglichkeiten für eine optimierte Aufgabengestaltung bei der computergestützten theoretischen Fahrerlaubnisprüfung. Zeitschrift für Verkehrssicherheit, 51 (3), 131-134.

Sturzbecher, D., Kammler, K., Weiße, B. & Breitling, K. (2009). Die Fahrerlaubnisprüfung im Spiegel der Bildungssoziologie. In J. Bönninger, K. Kammler & D. Sturzbecher (Hrsg.), *Die Geschichte der Fahrerlaubnisprüfung in Deutschland*. Bonn: Kirschbaum.

Sturzbecher, D., Kasper, D., Bönninger, J. & Rüdel, M. (2008). *Evaluation der Theoretischen Fahrerlaubnisprüfung – Methodische Konzeption und Ergebnisse des Revisionsprojekts*. Dresden: TÜV | DEKRA arge tp 21.

Sturzbecher, D., Mönch, M., Kissig, S.& Marschall, M. (2009). Die Entwicklung der Fahrerlaubnisprüfung in Deutschland von den Anfängen bis 1945. In D. Sturzbecher, J. Bönninger & K. Kammler (Hrsg.), *Die Geschichte der Fahrerlaubnisprüfung in Deutschland.* Dresden: TÜV | DEKRA arge tp 21.

Sturzbecher, D., Mörl, M. & Kaltenbaek, J. (2014). *Optimierung der Praktischen Fahrerlaubnisprüfung.* Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", M 243. Bremerhaven: Wirtschaftsverlag NW.

Sturzbecher, D. & Weiße, B. (2011). Möglichkeiten der Modellierung und Messung von Fahrkompetenz. In TÜV DEKRA arge tp 21 (Hrsg.), *Innovationsbericht zur Optimierung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2009/2010*. Dresden: TÜV DEKRA arge tp 21.

Sümer, N., Ünal, A.B. & Birdal, A. (2007) *Assessment of Hazard Perception Latencies Using Real Life and Animated Traffic Hazards: Comparison of Novice and Experienced Drivers*. Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, University of Iowa

Swoboda, G. (2001). *Das Kraftfahrt-Bundesamt und seine Vorgängerbehörden im Wandel der Zeit.* Flensburg.

Taylor, T., Masserang, K., Pradhan, A., Divekar, G., Samuel, S., Muttart, J., Pollatsek, A. & Fisher, D. (2011*). Long-term effects of hazard anticipation training on novice drivers measured on the open road.* Iowa: University of Iowa.

Tellegan, A. (1992). Personality traits: Issues of definition, evidence and assessment. *Psychology and Inquiry, 12*, 19-31.

Thomas, J.R., Gallgher, J., & Lowry, K. (2003). *Developing motor and sport expertise: Meta-analytic findings.* Paper presented at the North American Society for the Psychology of Sport and Physical Activity, 5-7 June, Savannah, GA.

Thomes, P., Kampker, A., Vallée, D., Schnettler, A. & Kasperk, G. (2013). Grundlagen. In A. Kampker, D. Vallée & A. Schnettler (Hrsg.), *Elektromobilität. Grundlagen einer Zukunftstechnologie* (pp. 5-58). Berlin, Heidelberg: Springer.

Triggs, T.J. & Regan, M.A. (1998). Development of a cognitive skills training product for novice drivers. In *The Proceedings of the 1998 Road Safety Research, Education and Enforcement Conference* (pp. 46-50). Wellington, New Zealand: Land Transport Authority.

TÜV | DEKRA arge tp 21 (2008). *Handbuch zum Fahrerlaubnisprüfungssystem (Theorie)*. Dresden: TÜV | DEKRA arge tp 21.

TÜV | DEKRA arge tp 21 (2011). (Hrsg.). *Innovationsbericht zur Optimierung der Theoretischen Fahrerlaubnisprüfung – Berichtszeitraum 2009/2010*. Dresden: TÜV | DEKRA arge tp 21.

TÜV | DEKRA arge tp 21 (2013). *Das Fahrerlaubnisprüfungssystem und seine Entwicklungspotenziale – Innovationsbericht 2009/2010. Innovationsbericht zur Optimierung der Fahrerlaubnisprüfung im Berichtszeitraum 2009/2010*. Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 239. Bremerhaven: Wirtschaftsverlag NW.

TÜV | DEKRA arge tp 21 (2015). Tätigkeitsbericht 2014 – Bericht zum Betrieb, zur Evaluation und zur Weiterentwicklung der Theoretischen Fahrerlaubnisprüfung. Dresden: TÜV | DEKRA arge tp 21.

Underwood, G. & Crundall, D. E. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics, 41 (4),* 448-458.

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics, 50 (8),* 1235-1249.

Underwood, G., Crundall, D., & Chapman, P. (2002). Selective searching while driving: the role of experience in hazard detection and general surveillance. *Ergonomics, 45*, 1-12.

Underwood, G., Ngai, A. & Underwood, J. (2013). Driving experience and situation awareness in hazard detection. *Safety Science, 56*, 29-35.

Underwood, G., Phelps, N., Wright, C., van Loon, E. & Galpin, A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Ophthalmic Physiological Optics (25)*, 346-356.

van Emmerik, I.J.H. (2004). The more you can get the better: mentoring constellations and intrinsic career outcomes. *Solidarity at Work Occasional Paper Series, 48,* 1-35.

van Merriënboer, J.J.G., Clark, R.E. & de Croock, M.B.M. (2002). Blueprints for complex learning: The 4C/ID*-model. *Educational Technology Research and Development, 50(2),* 39-64.

Velichkovsky, B.M., Rothert, A., Kopf, M., Dornhöfer, S.M. & Joos, M. (2002). Towards an Express-Diagnostics for Level of Processing and Hazard Perception. *Transportation Research Part F: Traffic Psychology and Behaviour, 5 (2)*, 145-156.

Verordnung über die Zulassung von Personen zum Straßenverkehr (Fahrerlaubnis-Verordnung – FeV). Accessed on 03.06.2015. URL: http://www.gesetze-im-internet.de/fev_2010/.

Vlakveld, W. (2011). *Hazard anticipation of young novice drivers. Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations*. SWOV, Leidschendam, Netherlands.

Vlakveld, W.P. (2014). A comparative study of two desktop hazard perception tasks suitable for mass testing in which scores are not based on response latencies. *Transportation Research Part F: Traffic Psychology and Behaviour, 22*, 218-231.

Vollrath, M. & Krems, J. (2011). *Verkehrspsychologie*. Stuttgart: Kohlhammer

Wallace, P., Haworth, N. & Regan, M. (2005). *Best methods for teaching hazard perception and responding by motorcyclists.* Report No. 236. Melbourne: Monash University Accident Research Centre.

Wallis, T.S. & Horswill, M.S. (2007). Using fuzzy signal detection theory to determine why experienced and trained drivers respond faster than novices in a hazard perception test. *Accident Analysis and Prevention, 39 (6),* 1177–1185.

Wedell-Hall, P. (2013). *Innovation in Assessment. Driving Standards Agency*. CIECA Workshop "How to Measure Hazard Perception?" 31 January - 1 February 2013, Helsinki.

Weise, G. (1975). *Psychologische Leistungstests*. Göttingen: Hogrefe.

Weiß, T., Bannert, M., Petzold, T. & Krems, J.F. (2009). *Einsatz von computergestützten Medien und Fahrsimulatoren in Fahrausbildung, Fahrerweiterbildung und Fahrerlaubnisprüfung*. Berichte der Bundesanstalt für Straßenwesen, Reihe "Mensch und Sicherheit", Heft M 202. Bremerhaven: Wirtschaftsverlag NW.

Wells, P., Tong, S., Sexton, B., Grayson, G., & Jones, E. (2008). *Cohort II: A Study of Learner and New Drivers.* Volume 1. Main Report. Road Safety Research Report No. 81. London: Department for Transport.

Wetton, M.A., Hill, A. & Horswill, M.S. (2011). The development and validation of a hazard perception test for use in driver licensing. *Accident Analysis and Prevention, 43*, 1759-1770.

Wetton, M.A., Hill, A. & Horswill, M.S. (2013). Are what happens next exercises and self-generated commentaries useful additions to hazard perception training for novice drivers? *Accident Analysis and Prevention, 54,* 57-66.

Whelan, M.I., Groeger, J.A., Senserrick, T.M. & Triggs, T.J. (2002). *Alternative methods of measuring hazard perception: Sensitivity to driving experience*. Road safety: Research, Policing and Education: conference proceedings (CD-ROM). Adelaide (South Australia). Accessed on 04.03.2014. URL: http://acrs.org.au/files/arsrpe/RS020055.PDF.

White, M. (2006). *Sorry, Mate, I Didn't See You: A Plausible Scientific Explanation*. South Australia Department of Transport.

Wickens, C.D. (2008). Situation Awareness: Review of Mica Endley´s 1995 Articles on Situation Awareness Theory and Measurement. *Human Factors, 50 (3)*, 397-403.

Wickens, C.D. & Hollands, J.G. (2000). *Engineering psychology and human performance* (3rd ed). Upper Saddle River, NJ: Prentice Hall.

Wickens, C.D. & McCarley, J.S. (2007). *Applied attention theory*. Boca Raton: Taylor & Francis.

Wiedenbeck, S. (1999). The use of icons and labels in an end user application program: an empirical study of learning and retention. *Behaviour & Information Technology, 18 (2),* 68-82.

Williamson, A.R. (2008). *Effect of video based road commentary training on the hazard perception skills of teenage novice drivers.* Waikato: The University of Waikato

Willmes-Lenz, G. (2010). *Unfallrisiko „Junge Fahrer" – neue Lösungsansätze.* Referat auf dem 48. Deutschen Verkehrssicherheitstag in Goslar, 27.-29.01.2008.

Winner, H. & Weitzel, A. (2012). Quo vadis, FAS? In H. Winner, S. Hakuli & G. Wolf (Hrsg.), *Handbuch Fahrerassistenzsysteme. Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (2., korrigierte Auflage). ATZ/MTZ-Fachbuch (pp. 658-667). Wiesbaden: Vieweg + Teubner.

Wolfe, J.M. & Horowitz, T.S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience 5,* 495-501.

Young, A.H., Chapman, P. & Crundall, D. (2014). Producing a commentary slows concurrent hazard perception responses. *Journal of Experimental Psychology, Applied, 20 (3),* 285-294.

ZEIT ONLINE GmbH (2014). Elektromobilität: Toyota will ab 2015 Brennstoffzellen-Autos verkaufen. Accessed on 05.08.2014. URL: http://www.zeit.de/mobilitaet/2014-06/toyota-brennstoffzellen-auto-elektromobilitaet.

Zimbardo, P.G. (1992). *Psychologie*. Heidelberg: Springer.

# Annex

# Task concepts for the assessment of abilities relating to traffic perception and hazard avoidance

## 1) Task concept "Distraction"

*Definition of the test content*

In a test item of this type, a dynamic, interactive traffic situation is presented. At first, the test person must handle only a single task drawn from the context of vehicle control (stabilisation, manoeuvring). This could, for example, involve use of the cursor keys to maintain an appropriate safe following distance behind a preceding vehicle. Subsequently, the test person is asked to perform a further additional task which is completely unrelated to the primary "driving task" and instead reflects a typical distraction in road traffic (e.g. entering a telephone number via a virtual keypad; entering a destination into a virtual navigation system). In this way, he is to be made aware of the risks which must be associated with distraction due to secondary activities while driving (e.g. use of a mobile telephone).

Comparable tasks addressing the issue of divided attention in road traffic were already developed at the Monash University Accident Research Centre in Australia in 2000 (Regan, Triggs & Godley, 2000). They were used firstly to demonstrate the effects of distraction to driving licence applicants, but at the same time also served to train attentional control as a basis for the handling of multiple cognitive demands. To this end, the set demands – the driving task and a mathematical question − were assigned different weighting over the course of the learning process, i.e. the person conducting the study specified which of the two tasks was to be given priority treatment.

The issue of distraction when participating in road traffic can be approached in a variety of ways with test items of this type, for example by modulating the demands in the initial "primary task" and the subsequent "secondary activity" (different road, traffic and weather conditions, different speeds, or different types and intensities of distraction).

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", "Localisation", "Identification", "Assessment of the hazard", "Assessment of own abilities", "Weighing up of subjective risk", "Decision" and "Action".

*Reference to learning objectives*

The fundamental learning objective of this task concept – namely development of an understanding for the negative safety impact of distraction – is presumably adequately clear to learner drivers. Feedback should focus predominantly on illustration and objective explanation of the drop in performance in the primary task associated with the driving process when attention is shared with a distracting secondary activity. Against this background, the concept seems suitable above all for use in the area of training. A training approach with the objective of improving the learner driver's ability to perform parallel tasks would presumably be counterproductive and is expressly not the intention of this task concept.

*Instruction and answer format*

The task instructions comprise textual components to explain the primary task (e.g. "Follow the vehicle ahead of you and maintain a safe following distance.") and to describe the secondary activity to be performed (e.g. "You have lost your way. Continue to follow the vehicle ahead of you at a safe distance. At the same time, enter the specified address in the navigation system."), alongside an interactive dynamic situation presentation (computer-generated). The required response is entered via the keyboard (e.g. cursor keys or the letters "S" and "F" to maintain the safe following distance; right-hand numerical keypad to enter the address or number). One possibility for variation of the task is to specify certain time limits for completion of the distracting activity.

*Performance parameters and assessment criteria*
Performance is measured by comparing how well a safe following distance is maintained, firstly as a single task and secondly as one part of a double task (e.g. percentages to specify the proportion of time during which the vehicle was within the correct distance range). The performance of the activity serving as a distraction must also be evaluated.

*Evaluation findings*
Experimental studies in a driving simulator confirm the effectiveness of this type of task as a means to improve attentional control. The effects were measured not only directly after the completion of training, but also several weeks later (Regan et al., 2000). No studies exist on the possible impact of such tasks on the understanding for hazards arising from distraction.

*Test economy and technical feasibility*
Prerequisite for the method is the availability of an interactive driving simulation with the possibility to record driving performance parameters relating to the primary task. A computer with monitor and mouse is required to actually perform the task (alternatively, a tablet with touchscreen could also be used).

**2) Task concept "Anticipation latency"**

*Definition of the test content*
In a test item of this type, the test person is shown a dynamic situation presentation, in which other road users (e.g. other vehicles, pedestrians, cyclists) are to be seen from the driver perspective. The task for the test person, while watching the video sequence, is to react with a computer input (mouse click, touchscreen, keyboard) as soon as he considers it necessary to brake in order to avoid an arising hazard. The intention is to determine whether novice drivers are able to detect arising hazards at an adequately early point in the development of a situation.

The anticipation of traffic situations is a prerequisite for timely behaviour adaptation and hazard prevention. Anticipation in such cases is based on corresponding information cues, which must be selected from a multitude of available stimuli and assessed with regard to their relevance. Experienced drivers apparently accomplish this better and faster than novice drivers. The performance advantage is often only of the order of a few milliseconds, but even that translates into significantly different reaction and stopping distances for proficient and less proficient drivers in real traffic.

This task concept is already used in Great Britain and Queensland as the basis for a traffic perception test within the framework of driving licence testing. The length of the time window for the required reaction input is determined on the basis of empirically measured performances by experts; it begins with the moment a hazard cue first becomes visible. In Queensland, the candidate is furthermore required to localise the hazard cue on the monitor screen; this serves to minimise the likelihood of random clicking being recorded as a correct response.

Test items of this type are presumably most suitable for content areas in which other road users are to be identified as manifest hazard cues, as it is there possible to present the development of potential conflicts in a vivid manner. Even so, various latent hazard cues (e.g. a partially concealed junction with a side road) are also conceivable as task content. It is furthermore possible to select entirely different contexts for the hazards to be anticipated (e.g. roads within and outside built-up areas, driving during the daytime and at night). Despite the aforementioned diversity of possibilities for content presentation, only a limited number of overarching demands can be expected overall, as the hazard cues to be detected are in the end always other road users (e.g. pedestrians between parked vehicles, vehicles approaching from the left or right, cyclists, vehicles pulling out of a parking space) with whom a collision may be imminent.

This task concept possesses references to the following demand components (see Chapter 2.2.3): “Observation”, possibly “Localisation” (e.g. if information visible in the vehicle mirrors is also relevant), “Identification” and “Assessment of the hazard”.

*Reference to learning objectives*
The learning objective of this task concept seems transparent: The fundamental importance of “timely hazard detection” should be obvious to novice drivers; this is also an expression of the principle of foresighted driving which is conveying during driver training. Especially where other vulnerable road users (e.g. a child) are to be detected, the practical relevance of the hazard cue should be readily understandable.

Learning-oriented feedback can be provided by presenting the video sequence a second time with appropriate markers to highlight the hazard cues which were to be detected. It is equally possible to continue presentation of the scenario beyond the critical point (up to which a reaction would have been able to avert an accident).

This task concept seems suitable both for training purposes (e.g. for learner assessments or in the form of exercises) and for the testing of novice drivers.

To facilitate acquisition of the necessary competences, McKenna and Crick (1994b) developed a four-hour training program which was realised primarily in a classroom group setting. Alongside the structured communication of knowledge on potential hazards, the focus was placed on enabling the participants to improve their application of visual search strategies and to anticipate the further course of an unfolding traffic situation. From further studies on the effectiveness of the specific teaching and learning methods applied in the training concept, the authors concluded that the observed training effects were attributable solely to the task of predicting the further course of a situation.

It can be assumed that this concept is suitable mainly for assessments of traffic perception involving other moving road users (pedestrians, other vehicles, etc.). Empirical testing is necessary to determine whether and, if so, to what extent content relating to hazards attributable above all to the driver's own (excessively) high driving speed (e.g. tight bends, stationary obstacles) are also suitable. Understanding for the intention of the task is possibly reduced where such content is addressed.

*Instruction and answer format*
The task instructions comprise a textual component (e.g. “Press the mouse button when you would apply the brakes”) and a dynamic situation presentation (e.g. a real video or computer-generated scenario). The required response is a mouse click when an action-relevant hazard cue is detected. If appropriate, the hazard cue concerned can then also be marked on the screen after a click.

Further possibilities for variation exist when designing test items of this type: The durations of the dynamic situation presentations can be generally shorter or longer. With increasing length of the depicted scenario, it is possible that performance differences in respect of vigilance may become significant; for the typical durations of less than one minute, however, this will probably be of no practical importance. Furthermore, a video sequence may contain and require the detection of more than one hazard, in which case attentiveness must be maintained over the whole duration. The hazards may be presented exclusively within the forward field of vision, though implementations involving the rear-view mirror are likewise feasible. The latter would presumably only be meaningful if the driver's own turning intentions are also signalled by way of a turn-indicator light on the dashboard (e.g. it could be visualised that a left turn is planned and at the same time a vehicle can be seen approaching from behind in the rear-view mirror).

*Performance parameters and assessment criteria*
The parameter used for performance assessment is the response latency, i.e. the delay between the moment a hazard cue first becomes visible and the clicking of the mouse to indicate that a hazard has been recognised. The reaction must fall within a defined time window. Assessments can be based on a multi-level rating scale dependent on the measured response latency; faster detection is treated as better performance.

It can be expected that classification on the basis of performance differences will only be successful with certain specific situational information cues (e.g. the cues should not be too obvious; they must emerge gradually).

*Evaluation findings*
Various studies have shown that this task concept is capable of distinguishing between novices and experts on the basis of performance (criterion: Driving experience) (e.g. Sexton, 2002; Grayson & Sexton, 2002; McKenna & Crick, 1994b).

*Test economy and technical feasibility*
Implementation of the method would be very expensive if based on specially produced real videos. The outlay is reduced if computer-generated sequences are used. Depending on the chosen length of each scenario, a certain minimum duration must be planned for test (e.g. 20 sequences of 60 seconds each), taking into account also additional time for practice items, instructions to the candidate and feedback. A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

### 3) Task concept "Case study"

*Definition of the test content*
This task category is conceived as a conglomerate of different question and item formats which together refer to a particular traffic situation or else a sequence of different situations in the sense of a "case study"; it thus bears similarity to a task concept used in the UK "Theory Test", where the candidate must answer several multiple-choice questions relating to a textually described traffic situation. The test items suggested here, however, are based on presentation of a dynamic visualisation viewed from the driver perspective. The scenario is halted at various points, whereupon a question is displayed and the test person is required to enter a corresponding answer. Possible question formats are: Multiple-choice questions (i.e. one of the specified answer options must be selected), gap-fill questions (i.e. the question must be answered by entering a number), free-text questions (i.e. an answer must be entered in the test person's own words), and assignment questions, here in the sense of tasks relating to the perspectives of other roads users (i.e. for a given object, the test person must mark the position of the object after the scenario is halted and the visualisation switches to a bird's-eye perspective). The intention of this concept is to determine whether novice drivers are able to judge traffic situations correctly from diverse perspectives, and to train them to pay due attention to the overall traffic environment.

The availability of different question formats means that case study tasks are suitable to portray both manifest and latent hazard cues. It is possible not only to present a multitude of qualitatively distinct task contents, but also to combine different item formats in accordance with the learning objectives to be addressed by a particular situation presentation.

The demand components (see Chapter 2.2.3) referenced by this task concept depend on the concrete situation presentations and the chosen item formats. Generally speaking, the following partial demands can be addressed: “Observation”, “Localisation”, “Identification”, “Assessment of the hazard”, “Assessment of own abilities” and “Weighing up of subjective risk”.

*Reference to learning objectives*
The specific learning objectives are determined by the nature of the chosen case study, the situation presentations used for visualisation, and the individual item formats and contents. The possibilities for feedback are dependent not least on the chosen item format. Essentially standardised feedback at the computer, for example, would appear to be especially feasible in conjunction with multiple-choice, gap-fill and assignment questions. To improve understanding of the feedback, explanation of the correct answers could be illustrated with static images taken from the originally dynamic situation presentation to depict the traffic situation at the moment at which the question was asked.

This concept seems suitable above all for use in the area of novice driver training.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. “Where is the motorcyclist?”) and a dynamic situation presentation (e.g. a real video or computer-generated scenario). The required response is a mouse click or keyboard input.

One possibility for variation of the task is to use static images instead of a video. These images could be displayed briefly and combined with one of the question formats. Another variation could be to limit the time allowed for the response.

*Performance parameters and assessment criteria*
Performance is judged according to the correctness of the selected answer option (for multiple-choice questions), the selected number (for gap-fill questions), the marked position (for assignment questions) or the given explanations (for free-text questions).

One limitation to be mentioned in this context is that deviations in the responses to assignment questions, i.e. indication of the positions of other road users, could be attributable to memory effects (and not, as intended, exclusively to the level of situation awareness). Furthermore, it is necessary to define a range within which the response to an assignment question will be accepted as correct. In the case of free-text questions, considerable outlay must be expected for evaluation of the replies.

*Evaluation findings*
No findings are available to date on the combination of different item formats within an overall case study. Multiple-choice questions are already in use in conjunction with dynamic situation presentations in the TDT and have there proved well suited as an instruction/answer format for the presentation of complex traffic situations and the response to correspondingly referenced test items.

*Test economy and technical feasibility*
Implementation of the method described here would be very expensive if the case studies are to be based on real videos. The outlay is reduced if computer-generated sequences are used. A computer with monitor and mouse is required (alternatively, a tablet with

touchscreen could also be used). For the most part, the responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary. In the case of free-text questions, on the other hand, evaluation must be entrusted to an appropriately trained person.

## 4) Task concept “Hazard classification and assessment”

*Definition of the test content*

The aim of this task concept is to determine situation awareness skills in complex traffic situations, placing a focus on hazard detection and overall assessment. It refers to the assessment of different types of hazard (latent, manifest, developing, etc.) as mobile or static elements of a traffic situation viewed from the ego perspective. To this end, the test person is shown a traffic situation as a dynamic visualisation (passive simulation). This scenario is then “frozen” at a certain point, with the final frame of the video remaining visible as a static image. Parallel to this, a menu of pictograms standing for different possible categories of hazard or hazard source is opened. The task is to specify whether a hazard is present and, if so, to select a maximum of two appropriate hazard categories. The basis for the actual selection and naming of hazards is thus a static image in which a specific hazardous traffic situation is depicted, characterised by certain elements of the traffic environment. In this way, it can be assessed whether the novice driver is able to classify and judge complex traffic situations in respect of their hazard potential.

The background to this task concept can be seen in study findings which indicate that novice drivers display deficits with regard to the early detection of hazards, and that this could well be attributable, at least in part, to inefficient gaze patterns. One source of such inadequate visual scanning behaviour may be the flawed situation schemata employed by novice drivers, i.e. they are less aware of the hazards which are especially likely to arise in a particular situation (and similarly of where to look for corresponding cues). The aim of test items of this type is thus not the localisation of perceived hazards (‘hotspotting’, Scialfa et al., 2011), but rather classification according to their content relevance. This should help to raise the novice driver's awareness for individual elements of the traffic space (pedestrians, cyclists) not per se, but rather in the context of a specific situation (Hole, 2007; Castro, 2009).

With the aid of a system of pictograms elaborated by experts (hazard catalogue), the intention is to accelerate the internalisation of selected hazards and hazardous situations, while at the same time encouraging awareness for specific potentially hazardous elements in the traffic space by way of abstraction. The pictograms are to be elaborated on the basis of lexical and semantic classifications (see Nakamura & Zeng-Treitler, 2012). The former refers to descriptions of entities as conceptual units or physical objects, whereas the latter describes events as processes or activities. The pictograms used should resemble actual road signs as closely as possible, and thus reflect the previously learned meta-language. The objective is an abstract and – as far as possible – visually unambiguous representation of hazards.

The advantage of pictograms is that they are easy to grasp visually and, in the ideal case, avoid the need for wordy explanations. In accordance with the semantic network theory (e.g. Collins & Loftus, 1975), stored information can be viewed as a network of interrelated concepts or nodes. The elaboration of a concept (e.g. a specific hazardous situation and the correct response to this hazard) serves to define relationships or links between different concepts or pieces of information (including that which has already been learned and stored). The result is a mutual strengthening of the associated and established concepts (see Lesch, 2008, p. 1006) and, for example, the facilitation of systematic hazard avoidance in the sense of conscious perception of the traffic space. Especially for novice drivers, after all, complex situations are difficult to grasp. The necessary prerequisite for early hazard detection and selection of an appropriate reaction, however, is an understanding of the overall situation; in this context, it is known, furthermore, that novices are slower in their reactions to perceived hazards compared to experienced drivers (Crundall et al., 1999; Deery, 1999; Sagberg & Bjørnskau, 2006), spatially distant hazards are perceived less successfully (Brown, 1982) and in part misinterpreted (Brown & Groeger, 1988), the peripheral vision of novice drivers is less well trained to draw information from the traffic space than that of experienced drivers (Mourant & Rockwell, 1972) and, according to Soliday (1974), static elements are better perceived as hazards than dynamic elements. Visual searching here means active searching for events and involved objects, even though it remains uncertain when, where and whether something will happen (e.g. Underwood, 2007).

The decision to use pictograms stems from the assumption that it is not sufficient to possess knowledge of potential hazards; more important is the ability to call up the necessary information at critical moments. Pictograms could be assigned a kind of "trigger function" within the framework of training, so as to enable information to be retrieved when needed (Lesch, 2008). In a test situation, such knowledge could be called up in a converse manner. By using abstract icons, it is possible to avoid providing solution hints.

The task concept described here requires the detection, assessment and reflection of hazards in a given traffic situation and also permits implicit hazards to be addressed. As in the case of the multiple-choice questions currently used in the TDT, there are extensive possibilities to create variants of a situation (superficial attributes of the traffic situations – e.g. driving environment, buildings, vehicle types and colours, vegetation, etc. – can be varied without affecting the thematic and demand-referenced core of the task).

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", "Identification" and "Assessment of the hazard".

*Reference to learning objectives*
The extent to which novice drivers recognise the learning objective associated with this task concept depends above all on the quality of the underlying system of pictograms and the implementation of this system in the actual items. As a necessary prerequisite for learning objective transparency, it is expedient to use a system of categories with proven reference to gaze patterns, search strategies and driving behaviour. Learning-oriented feedback could take the form of a brief explanation of the types of hazard presented (What constitutes a hazard and why?). Where the method is combined with a dynamic situation presentation, feedback could also be provided by allowing the scenario to continue beyond the stop marker, in order to reveal the subsequent development of the situation.

This concept seems suitable both for training purposes (e.g. for learner assessments or in the form of exercises) and for the testing of novice drivers. No empirical findings exist on the trainability of the abilities concerned, and no possible training programmes are known.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. "Which hazard(s) did you notice?"), a dynamic situation presentation (e.g. a computer-generated scenario) and a static situation presentation (e.g. a computer-generated image). The required response is entered by selecting a maximum of two hazard pictograms from the given options (e.g. by clicking with the mouse or touching the screen). Alongside the various hazard pictograms, a further pictogram is provided to denote "No hazard". The time allowed to enter a response should be limited to a duration which is still to be specified exactly.

A possible variation of the task could consist in implementation as a second level to a preceding hazard perception task, i.e. to follow up a task such as “Click with the mouse when you would apply the brakes to avoid a hazard” (anticipation latency). The static image to be judged would here be an interactive screenshot from the moment the mouse was clicked. If no click is entered and thus no hazard detected, an automatically generated screenshot can be presented at the end of the sequence; a corresponding pictogram “No hazard” can then be selected.

As a further variation, it is possible to forego the dynamic leader and to present only static images for evaluation and classification. With regard to the answer format, furthermore, the selection of a particular pictogram could be combined with the requirement to specify an explanation of this choice. The actual hazard would then have to be reflected and named. This could be implemented in the form that holding the mouse pointer over a pictogram displays three specific options to define the hazard more precisely (e.g. Hazard “Ice cream van” – Icon “Obstructed view (mobile)”: A) “A cyclist could overtake on the right”; B) “The lighting conditions could lead to dazzling after passing the vehicle”; C) “A child could run onto the road behind the vehicle.”).

*Performance parameters and assessment criteria*
Performance is measured according to whether the hazards defined for each image are identified correctly. The number of hazards in each case lies between 0 and 2. Driving on a wet road, for example, could be associated with both the concrete weather condition “Rain” and the corresponding hazards of reduced visibility and aquaplaning. If, in the course of the simulation, it starts to rain only lightly, there is no need to slow down due to a risk of aquaplaning; the reduced visibility, however, is very much a reason for deceleration. The total number of correct solutions is a conceivable performance index.

One limitation to be mentioned in this context is that the icons do not permit a specific hazard to be named. The test person must apply abstraction to the situations. A further difficulty lies in the clear distinction of evidently hazardous and evidently irrelevant aspects in a particular traffic situation.

*Evaluation findings*
No evaluation findings exist to date on the extent to which the ability to classify hazards according to categories such as “Weather”, “Pedestrians”, “View obstructed by static objects”, etc. correlates with improved performance in the detection of and reaction to hazards.

The use of pictograms (icons) may be helpful for learning processes, although the success is dependent on how users perceive and interpret the icons in each case. Misinterpretation may have a negative influence on understanding and could delay learning effects (Wiedenbeck, 1999). Pictograms with captions have been used in accident scenario training. Within the framework of this training, relationships were established between various warning symbols and corresponding accident scenarios (not necessarily in the road traffic context), and different aspects of knowledge relating to the symbols were then tested, e.g. verbal label, required/prohibited actions and possible consequences of the failure to comply. Training was found to result in significantly improved comprehension, shorter reaction times and an increased level of confidence in the responses given (Lesch, 2008).

*Test economy and technical feasibility*
The creation process would be relatively inexpensive if computer-generated images are used. A test duration of approx. 30 seconds per test item is to be planned.

A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

### 5) Task concept "Hazard localisation"

*Definition of the test content*
In a test item of this type, the test person is shown a static image of a traffic situation, with this image being divided into (possibly numbered) sectors. The task is to indicate the image sector where a hazard is to be seen. The intention is to determine whether the novice driver is able to recognise situation attributes with the potential to evolve into a hazard as the situation develops.

This type of task was devised by Hampel (1977a) as an alternative to classic multiple-choice questions. The new design demonstrated that it was possible to elaborate essentially non-verbal test items: It is not necessary to describe the hazard cues to be detected in verbally formulated answer options; instead, it is sufficient to offer the numbers of image sectors which may contain a potential hazard for selection.

From the thematic perspective, this concept is probably best suited for questions which address hazards in connection with other road users. It can be assumed that a hazard or hazard sector will only be identified correctly, if a hazard-relevant object which is sufficiently self-explanatory is actually to be seen in the image, e.g. a car door being opened, a vehicle reversing out of a parking space, a pedestrian. (Partially) concealed hazard cues, on the other hand, are presumably difficult to operationalise, precisely because they are not visible.

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", possibly "Localisation" (e.g. if information visible in the vehicle mirrors is also relevant), "Identification" and "Assessment of the hazard".

*Reference to learning objectives*
The learning objective of this task concept, namely the ability to recognise hazard cues, is presumably adequately clear to learner drivers. As feedback in case of an incorrect response, the correct sector of the image or the relevant object could be highlighted in an appropriate colour. It seems more difficult, however, to provide differentiated, learning-oriented feedback. The mere highlighting of the correct hazard cue or image sector conveys no information on the actual hazard to be expected. It is precisely this information which is relevant, however, as incorrect responses may also result from false assumptions regarding the significance of certain situation attributes. Differentiated feedback would thus presumably require an additional textual explanation. Where the method is combined with a dynamic situation presentation (see *Instruction and answer format* below), feedback could also be provided by allowing the scenario to continue beyond the point at which the response is entered, in order to reveal the subsequent development of the situation

This task concept seems suitable both for training purposes (e.g. for learner assessments or in the form of exercises) and for the testing of novice drivers. No empirical findings exist on the trainability of the demand components addressed.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. "Where is a hazard to be seen?") and a static situation presentation (e.g. a real photo or computer-generated image). The required response is entered by specifying a sector of the image (e.g. by way of the assigned number, by clicking with the mouse).

As an alternative to the above, the question could also be formulated in various other ways, for example “Where is it now important to look?”, “Where must you reckon with another vehicle?”, “Why must you wait before turning at the junction?” or “Why is it important to slow down here?”

Two possibilities for variation compared to the format originally suggested by Hampel (1977a) consist in the incorporation of a dynamic scenario as a leader to the final static image and furthermore limitation of the time allowed to enter a response. The use of a video sequence would presumably enhance the instruction quality, as the development of certain hazardous situations may already commence before the point at which a response is required (e.g. a parked vehicle can be seen in the distance, and when this point is “reached” in the final image, there is a risk that the driver could suddenly open the door). This possibility of anticipation would also justify the limitation of the permissible response time: If the developing situation has been followed attentively, and if all information cues have already been spotted in good time, it only remains to click on the appropriate image sector at the end of the scenario. It is thus no longer necessary to spend time searching for the hazard in the static final image.

*Performance parameters and assessment criteria*
Performance is judged according to whether the relevant hazard sector was selected. The hazard sectors can be understood as the answer options to a multiple-choice question and can be evaluated accordingly.

One limitation to be mentioned in this context is that the segmentation of the image (e.g. into five rectangular sectors) may lead to inaccuracies in performance assessment, firstly because a hazard-relevant object may not always fit perfectly into one sector, and secondly because an object may not fill the sector entirely (there could be uncertainty as to what exactly was considered relevant in a specified sector). If the sectors are defined to visibly match certain object contours (instead of standard rectangular sectors), on the other hand, this could effectively provide solution hints or else result in excessively differentiated sectors. To avoid this problem at least in part, it could be expedient to work with “invisible” image sectors (i.e. defined sectors of the image which serve merely to evaluate the position of the recorded mouse click).

*Evaluation findings*
Empirical findings relating to this task concept are available in a study conducted within the framework of the International Transport Exhibition in Hamburg in 1979 (Bönninger, Kammler & Sturzbecher, 2009). This study only involved two such test items, however. The results showed that participants with driving experience tended to solve the test items better than those without driving experience. Relatively speaking, the proposed test items attracted the least acceptance in a comparative analysis alongside other task formats (e.g. illustrated multiple-choice questions with very generalised answer options); in absolute terms, however, there were nevertheless around 65 per cent of the participants who considered this type of task to be suitable.

*Test economy and technical feasibility*
The required images could be generated relatively economically with the aid of a computer. Established evaluation procedures for multiple-choice questions could be taken over (if necessary with only minor adaptation). The time to be planned for completion of the tasks corresponds to that known from tests comprising multiple-choice questions. A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive visualisation of a traffic situation, i.e. an interactive presentation is not necessary.

### 6) Task concept “Action selection”

*Definition of the test content*

In a test item of this type, the test person is shown a static image of a traffic situation. Within a given response time of just a few seconds, the action which is appropriate in the depicted situation must be selected from the three action options offered. This task concept is already used in the Netherlands as the basis for a traffic perception test within the framework of driving licence testing. The test there comprises 25 photographs of traffic situations viewed from the driver perspective (each including relevant information in the mirrors, and with the turn indicators and speedometer visible). The candidate is allowed up to eight seconds to decide which of the three possible behaviour options “Apply the brakes”, “Take foot off the accelerator” or “Do nothing” is appropriate in the depicted situation; these same options are presented in all of the 25 test items. The intention is to determine whether the test candidate is able to reach appropriate action decisions, taking into account the circumstances of a specific traffic situation.

With regard to the theoretical action demand, the operationalisation implemented in the Netherlands refers to the component “Action selection” in the model elaborated by Grayson et al. (2003). The three action options correspond to two different forms of depicted hazards: So-called “acute hazards” (e.g. a pedestrian who wants to catch a bus waiting on the other side of the road and crosses without paying attention to the traffic) require the reaction “Apply the brakes”; the reaction “Take foot off the accelerator”, on the other hand, would be appropriate in the case of “latent hazard” (e.g. a bus standing at a bus stop, and thus the possibility that passengers could alight and wish cross the road, even though no persons are actually visible in the depicted scene).

Since it is possible to address both latent and manifest hazards, such test items are generally suitable for a broad spectrum of demand situations (e.g. other road users, relevant attributes of the road, the negotiation of tight bends).

This task concept possesses references to the following demand components (see Chapter 2.2.3): “Observation”, possibly “Localisation” (e.g. if information visible in the vehicle mirrors is also relevant), “Identification”, “Assessment of the hazard” and “Decision”.

*Reference to learning objectives*

The fundamental learning objective of this task concept should be transparent for driving licence applicants, in that it refers to the identification of information cues for particular actions in a real traffic situation. It cannot be judged, however, whether the test persons actually understand why only one combination of situation and action is deemed acceptable. There is possibly a certain lack of clarity, especially with regard to the distinction between “Take foot off the accelerator” and “Apply the brakes”. Feedback would need to be correspondingly differentiated: The relevant hazard cue could be highlighted visually. At the same time, however, it would be important to explain why, for example, the option “Take foot off the accelerator” is deemed the correct reaction to a properly recognised hazard cue, but “Apply the brakes” is considered incorrect.

This concept seems suitable both for training purposes (e.g. for learner assessments or in the form of exercises) and for the testing of novice drivers. No information is available on specific training programs or studies relating to the trainability of the abilities concerned.

*Instruction and answer format*

The item instructions comprise a textual component (e.g. “Select the action which you consider appropriate in the depicted situation”) and a static situation presentation (e.g. a real photo or computer-generated image). In a Dutch test centre, the images are displayed on a single screen for a whole group of test candidates. The response of each individual candidate is recorded by way of a "voting device" (with the options A, B and C); various adaptations are conceivable, however, such as the use of marked keys on a standard computer keyboard.

A possible variation of the task – taking up certain methodical reservations (see *Performance parameters and assessment criteria* below) – could be limited to assessment of the depicted situation, without specification of an appropriate action. The static image could also be preceded by a dynamic scenario to illustrate the development of the situation. Accordingly, the test person could be asked to specify the nature of the hazard cue, if any, present in a given situation (i.e. none, latent hazard cue or manifest hazard cue). Another possibility could be a two-stage method, wherein it is first necessary to decide whether a depicted situation is potentially hazardous. If so, the next step would then be to click on the image in the area where the (latent or manifest) hazard is seen. It would here be beneficial to have internalised the concept of differently perceivable hazards.

*Performance parameters and assessment criteria*
The responses can be evaluated in the same way as a classic multiple-choice test. In the Netherlands, the candidate must react within a certain time window (of up to eight seconds), and at least 12 of the 25 test items must be completed correctly.

One methodical reservation to be mentioned with regard to the practical example from the Netherlands is that the answer options distinguish three actions which can be assigned precisely to three categories of situation (i.e. "Do nothing", if no hazard is present; "Take foot off the accelerator", if a latent hazard is present; "Apply the brakes", if a manifest hazard is present). Theoretically, at least, the test candidate must thus first decide which situation category applies, and then which action is appropriate. This reduces the nine theoretically possible permutations of situation and action to just three, i.e. the probability of merely guessing a correct response to a particular test item is 1:3. If a strictly logical approach is applied, the additional specification of different action options is in the end meaningless. In the least favourable case, candidates are even "penalised" if they fail to bear in mind (or understand) the logic of this construction, i.e. if they recognise a situation correctly, but react more cautiously than necessary (e.g. by selecting "Apply the brakes" or "Take foot off the accelerator" even though no hazard is present, or else "Apply the brakes" in case of a latent hazard).

*Evaluation findings*
Empirical studies conducted by Vlakveld (2011) revealed that driving experience had a significant influence on test performance: "Expert" drivers achieved higher overall scores than novices. In the expert group (comprising drivers with 18 months of driving experience), an additional analysis was performed to consider a possible correlation between the overall score achieved and the number of self-reported accidents, for which purpose differences in the amount of driving done were similarly acquired and taken into account in the evaluation as a factor which could influence accident frequency. The result of this analysis was that, on average, drivers with an accident-free driving record achieved significantly better test results than drivers who reported accident involvement. The findings reported by Vlakfeld (ibid.) suggest that traffic-relevant performance differences can be measured. As the driver groups were compared solely on the basis of their overall scores, however, it is not evident whether (and, if so, what) differences exist with regard to the individual demand aspects (i.e. the differentiation of latent hazards, manifest hazards and non-hazardous situations).

*Test economy and technical feasibility*
The creation process for computer-generated images would be relatively inexpensive. The evaluation procedures would also correspond to the established mechanisms applied

to multiple-choice tests. The test duration would presumably be similar to that which must be planned today for a corresponding number of multiple-choice questions. The technical implementation requires a computer with monitor and mouse (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

**7) Task concept "Action timing"**

*Definition of the test content*

In a test item of this type, the test person is shown various traffic situations in the form of dynamic scenarios. Before each video sequence (with a duration of, for example, 30 seconds), text instructions are displayed to specify the action which is to be performed (e.g. slow down, overtake or cross/turn at the junction). The task is to indicate by pressing a key, when (i.e. at which point during the course of the viewed scenario) the required action can be performed safely. Performance is rated according to whether the person reacts within a defined time window. A test series also includes items in which the specified action is never appropriate ("no response items").

The objective of this task concept is to assess whether novice drivers are able to derive the timing for safe performance of a required action from the given situation attributes. Such items are already used in the described and similar forms in the Australian states of Victoria, New South Wales, South Australia und Western Australia as the basis for a traffic perception test within the framework of driving licence testing; these tests are implemented with real-life video clips.

The test contents can be deemed empirically founded, as they are derived from the most frequent and most serious types of accident involving novice drivers. This concept seems especially interesting because it places subjective assessment of a situation in the foreground, i.e. it may also be able to distinguish risk-seekers and more cautious drivers.

In the aforementioned Australian states, the specified actions refer overall to the driving demands "Overtaking", "Turning at a junction", "Slowing down", "Crossing an intersection" and "Moving off". The use of different actions in a single test permits the construction of a multitude of different test items. As in the case of the multiple-choice questions currently used in the TDT, there are extensive possibilities to create variants of a situation (superficial attributes of the traffic situations – e.g. driving environment, buildings, vehicle types and colours, vegetation, etc. – can be varied without affecting the thematic and demand-referenced core of the task).

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", "Localisation", "Identification", "Assessment of the hazard", "Assessment of own abilities" and "Weighing up of subjective risk".

*Reference to learning objectives*

The learning objective of this task concept is presumably quite clear to novice drivers: The requirement to reach an action decision (e.g. turning at a junction) is plausibly related to real traffic demands. Learning-oriented feedback could be given by presenting the video sequence a second time in case of an incorrect response, for example with appropriate highlighting of the objects which are relevant for the action decision (e.g. an oncoming vehicle). The feedback should also reveal the times during the video sequence at which an action could be considered safe, and when it would be unsafe or dangerous; the status "safe" could be indicated through appropriate use of the colour green (colouring of the object itself, a progress bar at the top of the screen, or the like), while the colour red could signal the transition to the status "unsafe".

This task concept seems suitable both for training purposes (e.g. for learner assessments or in the form of exercises) and for the testing of novice drivers. No information is available on specific training programs or studies relating to the trainability of the abilities concerned. The applicable driver training systems refer to on-road driving practice as a means to prepare for the test.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. "You wish to turn left at the junction. Press the space bar when it is possible to turn safely.") and a subsequent dynamic situation presentation (e.g. a real video or computer-generated scenario). The point at which it is considered safe to perform the action is specified via a keyboard or by clicking with the mouse. The chosen moment represents the intended beginning for the driving manoeuvre specified in the instructions.

Various possibilities for variation exist when designing test items of this type. For example, relevant information cues could appear either exclusively within the forward field of vision or else additionally in the side/rear-view mirrors (e.g. the decision to turn at a junction or move off from the kerbside could be influenced by longitudinal traffic approaching from ahead and/or from behind). In another variation, the computer input could serve not to designate the point at which an action can be performed, but rather to abort a pending (unsafe) action. This reversal of the task would place fast prediction of the further development of a situation in the foreground (e.g. in a situation viewed from the driver's perspective, the test person's own vehicle signals the intention to overtake by setting the turn indicators and pulls out to do so, even though oncoming traffic is to be seen in the nearer distance; in this case, the mouse must be clicked to "intervene" as quickly as possible and to halt the scenario).

*Performance parameters and assessment criteria*
Performance is measured according to whether the action is appropriate at the time specified. The input must be recorded within a certain time window, i.e. the assessment is a dichotomous rating ("within" or "outside" the time window defined for safe performance of the action).

One limitation to be mentioned in this context is that situation construction seems to be very demanding: If situations which permit unambiguous judgements of "correct" or "incorrect" are constructed (e.g. the vehicle is standing at a junction and wishes to turn left; oncoming traffic is approaching and it is thus not possible to make the turn), they will probably fail to discriminate. If more complex situations are chosen, however, it will probably be more difficult to elaborate sound foundations for assessment of the displayed performance from the perspective of test didactics (e.g. if the non-performance of an action is treated as an error because it would not have been dangerous). It should also be taken into account that a dichotomous classification of "safe" and "unsafe" situations could also detract from the validity of the performance assessment: Hazardous situations in which an action should not be performed must always be relatively unambiguous (otherwise, performance would have been possible). If the oncoming traffic is still in the far distance when turning left at a junction, the action must be considered safe. If an oncoming vehicle is already close enough to constitute a hazard, on the other hand, it would presumably also be obvious that the turning manoeuvre should no longer be performed.

*Evaluation findings*
Evaluation findings presented by Congdon and Cavallo (1999) seem to confirm the predictive validity of this type of task. Persons with a relatively poor test result were up to three times more likely to be involved in a serious or fatal accident than persons with relatively good results. Using a further developed version of the same test method, it

could be shown that better test performances tended to correlate with greater driving experience (Catchpole, Congdon & Leadbeatter, 2001).

*Test economy and technical feasibility*
The development process would be very expensive if based on specially produced real videos. The outlay could be reduced if computer-generated sequences are used. The test duration to be planned is essentially dependent on the length of the individual scenarios (e.g. 30 seconds each), and can hardly be longer than the time required for the dynamic situation presentations already in use today (plus additional time for practice items, instructions to the candidate and feedback).

A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

**8) Task concept "Object assessment"**

*Definition of the test content*
In a test item of this type, the test person is shown a static image of a traffic situation for a few seconds (a displayed timer counts down: "3", "2", "1"). When the timer expires, the image disappears for a moment, and then returns in slightly modified form. In the second image, three objects have been changed or added (e.g. tree, vehicle or pedestrian), one of which is hazard-relevant (e.g. pedestrian standing at the side of the road). The test person must click on the hazard-relevant object within a defined time window; the other two objects serve as distractors. The intention is to assess how quickly novice drivers are able to classify objects in a concrete traffic situation as hazard-relevant or irrelevant.

Studies conducted by Wetton et al. (2011) show that novice drivers are no less able to detect changes in a situation than experienced drivers; in fact, their test reactions were actually faster. It is possible, however, that the deficits of novice drivers are manifested not on the elementary level of object recognition and change detection, but rather in terms of their ability to assess whether an object is hazard-relevant. The proposed task format expands the perceptual psychology research paradigm of "change detection" in the sense that it demands not only the detection of modified objects, but also their classification and assessment.

Such test items are suitable above all for content relating to other road users, as the latter can here be represented directly as hazard-relevant objects (vehicles, persons). A multitude of qualitatively distinct task contents can be created, as a broad diversity of hazard-relevant objects (pedestrians, cyclists, children, other vehicles) can be incorporated and the positions of these objects in the image can also be varied. As in the case of the multiple-choice questions currently used in the TDT, there are extensive possibilities to create variants of a situation (superficial attributes of the traffic situations – e.g. driving environment, buildings, vehicle types and colours, vegetation, etc. – can be varied without affecting the thematic and demand-referenced core of the task).

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", possibly "Localisation" (e.g. if information visible in the vehicle mirrors is also relevant), "Identification", "Assessment of the hazard".

*Reference to learning objectives*
The transparency of the learning objective is to a large degree dependent on the rationale behind assessment of an object as hazard-relevant. It is necessary to find unambiguous attractors and distractors which are at the same time self-explanatory for the test person. Feedback can be given by marking/highlighting the hazard-relevant objects.

This concept seems suitable above all for use in the area of training (e.g. for learner assessments or in the form of exercises), but could also be used for test purposes subject to certain limitations.

No empirical findings exist on the trainability of the abilities concerned, and no information is available on possible training programs.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. “Three objects have been modified or added in the depicted traffic situation. Click on the object which requires you to slow down in this situation.”) and a static image (e.g. computer-generated). The relevant object is selected by clicking with the mouse (or touching the screen).

Various possibilities for variation exist when designing test items of this type. For example, relevant information cues could appear either exclusively within the forward field of vision or else additionally in the side/rear-view mirrors. The extent of the modifications to individual objects can also be varied. It is furthermore essential for the construction process to consider carefully why a certain object should be treated as hazard-relevant, but the other two not. One relatively simple construction model could consist in always adding distractors in the form of a stationary object (e.g. parked car, tree, house). Another straightforward approach, as an alternative to the addition of distractors, could be to exchange objects or object attributes (e.g. vehicle model or colour). Such creation principles are economical, but could also provide pointers for schematic recognition of the required solution (e.g. “Always click the ‘moving’ object”, “Always click the object which was added”).

*Performance parameters and assessment criteria*
Performance is judged according to the total number of correct classifications of hazard-relevant objects. The three modified objects can be understood as three answer options to a classic multiple-choice question, one of which is correct. The task is solved if the relevant object is selected. The response must be entered within a defined time window.

One limitation to be mentioned in this context is that, due to the necessity of selection with the mouse, this type of task is suitable primarily for the classification and assessment of concrete objects (i.e. “explicit” hazard cues). It is true that it would be possible to modify the course of the road or the weather conditions, for example, but this would then affect a large section or even the whole image; that could result is mistaken responses. There are many traffic situations where the starting point is always one particular object with hazard relevance (e.g. vehicle approaching from the right, cyclist, child on the side of the road). It should be taken into account that such objects are probably easy to recognise as changed and hazard-relevant if they are placed directly in front of the test person's virtual vehicle. If objects are incorporated into the scene at a greater distance, it is more difficult to detect the modification, although they are then also less of an actual hazard. A further point to be noted is that it is not only the swiftness of classification (where novice drivers are presumably slower than experienced drivers) which is decisive for the test result; before any assessment is possible, it is first of all necessary to identify the three changes in the image (where empirical findings suggest that novice drivers may actually be faster, see above).

*Evaluation findings*
As this is a newly devised task concept, no evaluation findings exist to date.

*Test economy and technical feasibility*
The creation process would be relatively inexpensive, irrespective of whether real photos or computer-generated static images are used, though the modification of individual ob-

jects would be especially simple in the case of computer-generated images. The test duration would be quite short, as the images for each test item (before/after) must only be displayed for a few seconds (plus additional time for practice items, instructions to the candidate and feedback).

A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

**9) Task concept "Situation assessment"**

*Definition of the test content*
In a test item of this type, the test person is shown two practically identical images of traffic situations. Within the limited time permitted, he must then decide which situation is more hazardous (e.g. tight or sweeping bend; bend with clear or obstructed view of the road ahead; bend in which the vehicle is driving closer to the centre line or to the kerbside; wet or dry road surface; paved roadside or soft verge). One of the images always depicts a greater hazard than the other, as one of the situation attributes is modified to enhance the underlying risk. The intention is to determine whether novice drivers are able to derive the hazard potential of different traffic situations from implicit hazard cues.

The method is related to both the "Adaptation Test" developed by DeCraen (2010; see *Evaluation findings* below) – in that discreetly modified traffic situations must be compared – and the Dutch traffic perception test (see Task concept 6 "Action selection" above) – in that implicit hazard cues must be taken into account. In the Adaptation Test, the difference between the two images (e.g. a bend to the left on a road outside built-up areas, where the view of the road ahead is either unobstructed or else restricted by bushes) contributes to a greater or lesser complexity of the situation. The task for the test person is in each case to specify the driving speed which is considered appropriate in the depicted situation. A response is deemed correct if the speed selected for the more complex situation is lower than for the less complex situation (an identical speed for both images of a pair or a higher speed in the more complex situation are marked as incorrect). According to DeCraen (ibid.), the objective of the Adaptation Test is not merely to assess the detection of hazards; it goes further by asking for behaviour appropriate to the viewed situation. It should be kept in mind, however, that the demands of the Adaptation Test are higher than those of the described item format, as the situations are not presented simultaneously for direct comparison; it is instead necessary to make an absolute judgement on individual situations presented in a random order.

Such test items are presumably suitable above all for content relating to traffic infrastructure (e.g. road surfaces, bend radii), as they also enable implicit hazards to be depicted. The construction principle permits the creation of an adequate number of task variants in which it is only necessary for each depictable hazard (e.g. a bend) to be "intensified" (attractor) or "alleviated" (distractor) in respect of the relevant situation attribute. It is furthermore relatively simple to develop variants of an illustration.

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", "Identification" and "Assessment of the hazard".

*Reference to learning objectives*
The transparency of the learning objective is to a large degree dependent on the rationale behind classification of a response as correct or incorrect for a given situation comparison (e.g. due to the tightness of the bend, which entails greater hazard potential). Learning-oriented feedback must thus offer adequate explanations of the hazards addressed, espe-

cially where implicit hazards are involved. Potential hazards in connection with a soft verge or a bend with limited visibility, for example, are not necessarily self-explanatory.

This concept seems suitable both for training purposes (e.g. for learner assessments or in the form of exercises) and for the testing of novice drivers.

No empirical findings exist on the trainability of the abilities concerned, and no information is available on possible training programs.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. "Which of the two situations is potentially more hazardous for you as a driver? Compare the images and select the more hazardous situation.") and two static situation presentations (e.g. computer-generated images). The required response is entered by clicking with the mouse on the image which is considered more hazardous. The time allowed to enter a response is limited.

Various possibilities for variation exist when designing test items of this type. It could be meaningful, for example, to add a second question to each image pair to ask for an explanation of the previous decision. This would help to reduce the high probability of correct guessing. The question as to why the situation in the selected image is more hazardous could be accompanied by three verbally formulated answer options.

*Performance parameters and assessment criteria*
For performance assessment, the two images are treated as the answer options to a multiple-choice question. The task is solved if the image depicting the more hazardous situation is selected. The response must be entered within a defined time window.

The task concept "Situation assessment" seems appropriate above all for situations with latent hazard cues, e.g. situations relating to the risk of single-vehicle accidents in connection with tight bends. Where manifest hazard cues are involved, there is a risk of constructing image pairs where "change detection" is sufficient to determine the correct answer (e.g. "I must click on the image in which there is an additional person or vehicle, in which the person or vehicle is closer, and so on."). With latent hazard cues, on the other hand, it is necessary to assess the detected change in the image, because it is no longer sufficient merely to determine which image contains "more" (e.g. in an image pair where a tree-lined country road is shown with and without crash barriers, the image depicting the greater hazard is actually the one in which there are fewer elements).

*Evaluation findings*
As this is a new task concept, no evaluation findings exist to date. The results of the studies conducted by DeCraen (2010) with the "Adaptation Test" (see above), however, reveal that the average experienced driver gave correct answers more frequently than the typical novice driver. Similarly, drivers who overestimated their driving skills in the self-assessment, and likewise those who were identified as unsafe drivers within the framework of an on-road driving assessment, achieved poorer results.

*Test economy and technical feasibility*
Implementation of this method would be relatively inexpensive if computer-generated images are used. The test duration would be quite short, as each image pair must only be displayed for a few seconds (plus additional time for practice items, instructions to the candidate and feedback).

A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

## 10) Task concept "Situation awareness"

*Definition of the test content*

In a test item of this type, the test person is shown a short video sequence of a particular traffic situation. The video sequence is halted at the end and the final image also disappears after a few seconds. Subsequently, three further static images are displayed: Each image represents a different possible continuation of the scenario. The task is to select the future state which is deemed most probable in the given circumstances. In this way, it can be determined whether novice drivers are able to predict the short-term further development of the system "road traffic" on the basis of situational information cues.

This task concept is related to approaches for the measurement of situation awareness, as described in the scientific literature (see, among others, Malone et al., 2012). One example is the Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1995), where the test person tackles a given task in a simulator. The simulation is halted ("frozen") at certain points and the display is blacked out; questions relating to the previously viewed situation must then be answered (this is a so-called "objective" method, which compares knowledge with the actual state of a system; a subjective method, by contrast, relies on the use of rating scales or else third-party observations). The OSCAR model (Outil Standardisé pour la Comparaison et l'Analyse des Représentations mentales; Bailly, Bellet & Goupil, 2003) assesses whether the test person has developed an adequate mental representation of the situation by showing first a short video sequence and then a slightly modified depiction of the situation (e.g. traffic lights at a previously uncontrolled junction); the test person is required to click on the area of the image which has changed.

Such test items are presumably suitable above all for content relating to other road users, as predictions of object constellations and relations are required. A multitude of qualitatively distinct task contents can be created, as a broad diversity of hazard-relevant objects (pedestrians, cyclists, children, other vehicles) can be incorporated and the positions of these objects in the image can also be varied. As in the case of the multiple-choice questions currently used in the TDT, there are extensive possibilities to create variants of a situation (superficial attributes of the traffic situations – e.g. driving environment, buildings, vehicle types and colours, vegetation, etc. – can be varied without affecting the thematic and demand-referenced core of the task).

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", "Localisation" and "Identification".

*Reference to learning objectives*

The learning objective of this task concept should be transparent for driving licence applicants, as the "prediction" of situation developments is a clear training goal ("foresighted driving"). The fact that the situation to be predicted need not necessarily be hazard-relevant could reduce understanding for the learning objective. Learning-oriented feedback could be provided by showing the section of the scenario which was "skipped" when presented for the test item, i.e. by revealing how the situation at the end of the dynamic sequence evolved up to the point of the correct final image.

The concept seems suitable above all for use in the area of training (e.g. within the framework of exercises). If it is to be used for testing, the content would probably need to refer to hazard-relevant scenarios, though this would in turn facilitate the use of solution strategies (see *Performance parameters and assessment criteria* below).

No empirical findings exist on the trainability of the abilities concerned, and no information is available on possible training programs.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. before the video: “Watch the following video carefully. It will be faded out after 15 seconds, and you must then predict the further development of the situation”; after the video: “How did the situation develop from this point? Select the appropriate image.”), a dynamic situation presentation (e.g. a real video or computer-generated scenario) and static images (real photos or computer-generated). The correct image is selected by clicking with the mouse (or touching the screen).

One possibility for variation when designing test items of this type is that relevant information cues could appear either exclusively within the forward field of vision or else additionally in the side/rear-view mirrors.

*Performance parameters and assessment criteria*
Performance is judged according to the number of correct situation predictions. With three answer options, the probability of guessing the correct answer is 1:3. The situation images can be understood as the answer options to a classic multiple-choice question. The task is solved if the correct image is selected. The time allowed to enter a response should be limited, so as to prevent deduction of the correct solution from a close comparison of the three images (e.g. the possible discovery of inconsistent image elements in the distractors).

One limitation to be mentioned in this context is that the method is possibly best suited for the assessment of situation awareness in general, but less so for measurement of an awareness for specific hazards. In an application which addresses the relatively confined aspect of “hazard awareness”, even a simple heuristic (“Always choose the most hazardous situation”) would probably identify the correct solution to the majority of test items. This effect could perhaps be reduced if the objects serving as hazard cues (e.g. a cyclist) are also present in the distractors, but there in an “inconsistent” position. The gap in time between the video and the situation images offered for selection must be kept short, i.e. the situation must only be allowed to develop to a minor extent, as otherwise even the distractors could – at least hypothetically – become applicable. It is furthermore to be taken into account that this type of task places additional demands on the test person (memory performance; possibly explicit verbalisation of the visual information) which are not relevant in this form for natural traffic behaviour.

*Evaluation findings*
As this is a new task concept, no evaluation findings exist to date.

*Test economy and technical feasibility*
The outlay for implementation of this method can be reduced if computer-generated images are used, rather than real videos and photos. Computer-generated material also seems expedient since the answer options can all be derived from the same video sequence. The test duration to be planned is essentially dependent on the length of the individual video sequences. Assuming that an individual sequence lasts 15 seconds, it would hardly be longer than the time required for the dynamic situation presentations already in use today (plus additional time for practice items, instructions to the candidate and feedback).

A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, i.e. an interactive presentation is not necessary.

## 11) Task concept "Blind spot"

*Definition of the test content*
In a test item of this type, the test person is shown a dynamic scenario viewed from the driver's perspective. As soon as the scenario is halted, he must answer a multiple-choice question relating to the depicted situation. To be able to answer the question correctly, it is necessary to be aware of the traffic situation in the "blind spot". This intention of such tasks is to assess (and convey) knowledge relating to hazards in connection with a blind spot, alongside specific visual search strategies for hazard avoidance.

The background to this concept lies in scientific findings on the manifold deficits displayed by novice drivers in respect of glance behaviour and scanning techniques. One prominent example is the fact that novice drivers only rarely use the side and rear-view mirrors for information acquisition. Furthermore, they only seldom use shoulder checks to detect possible hazards in the blind spot (Underwood, Crundal & Chapman, 2002).

From the thematic perspective, such test items are limited to hazards in connection with a blind spot. There is presumably little scope to develop task variations: It is likely that certain overarching demand patterns (e.g. overtaking vehicles) will emerge, and that variation is only possible by modifying the environment in which the hazard cue is placed.

This task concept possesses references to the following demand components (see Chapter 2.2.3): "Observation", "Localisation", "Identification", "Assessment of the hazard", "Assessment of own abilities" and "Weighing up of subjective risk".

*Reference to learning objectives*
The learning objective of this task concept, namely awareness for the hazards associated with the blind spot and possibilities to avoid such hazards, is presumably adequately clear to learner drivers. "Virtual shoulder checks" could be used to provide learning-oriented feedback: After answering the multiple-choice question, the test person could shift the static final image to the left or right, to simulate the effect of turning his head to perform a shoulder check.

The concept seems suitable above all for use in the training of novice drivers (e.g. for learner assessments or in the form of exercises). The scope of application is too limited for use in testing. No empirical findings exist on the trainability of the abilities concerned, and no information is available on possible training programs.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. "Is it safe to switch to the left-hand lane at this point?") and a dynamic situation presentation (e.g. a computer-generated scenario). The required response is entered by clicking with the mouse on the defined answers.

*Performance parameters and assessment criteria*
Performance could be judged according to the total number of correct answers.

*Evaluation findings*
As this is a new task concept, no evaluation findings exist to date.

*Test economy and technical feasibility*
A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium; implementation of the proposed feedback, however, requires a certain degree of interactive functionality.

## 12) Task concept "Sweep to reveal"

*Definition of the test content*
In a test item of this type, the test person is shown a static image of the situation to which a question refers. Correct answering of the question is made difficult by obstructing the test person's view of the relevant object or person. To solve the task, it is necessary to "sweep aside" the obstruction by clicking with the mouse; the concealed hazard is then plainly visible. Once the object obstructing visibility has been removed, the question can be answered. The intention of this task concept is to show that hazards may be concealed behind other objects or persons, and are thus not always directly observable.

The background to the concept lies in scientific findings on the deficits displayed by novice drivers in the detection of implicit or concealed hazards. A similar item format is already in use in the learning aids published by Verlag Heinrich Vogel.

From the thematic perspective, such test items are limited to situations with concealed hazard cues. Correspondingly, the scope to develop multiple item variations is also limited: It is likely that certain overarching demand patterns (e.g. children behind a vehicle; junction hidden by bushes) will emerge, and that variation is only possible by modifying the environment in which the hazard cue is placed.

The demand components (see Chapter 2.2.3) referenced by this task concept depend on the questions asked. Generally speaking, the following demand components can be addressed: "Observation", "Localisation" and "Identification".

*Reference to learning objectives*
The learning objective of this task concept, namely the detection of concealed hazards, is presumably adequately clear to learner drivers. Learning-oriented feedback could be provided by marking the correct answer option. Additionally, the test person could be shown the original image (with obstructed view) alongside the corresponding "solution" version (without obstruction).

The concept seems suitable above all for use in the training of novice drivers (e.g. within the framework of exercises). No empirical findings exist on the trainability of the abilities concerned, and no information is available on possible training programs.

*Instruction and answer format*
The item instructions comprise a textual component (e.g. "Which element of this scene requires your particular attention?") and a static image (computer-generated) with corresponding interactive functionality. The required response is entered by clicking and sweeping with the mouse ("drag and drop"-style action).

The task could be varied by presenting a dynamic scenario as a leader to the final static image. The use of a dynamic leader would presumably enhance the instruction quality, as certain hazardous situations may already begin to develop in advance and could then be identified more effectively.

*Performance parameters and assessment criteria*
Performance is judged according to whether the correct answer option is chosen in response to the multiple-choice question.

*Evaluation findings*
No evaluation findings exist to date.

*Test economy and technical feasibility*
The creation process for computer-generated images would be relatively inexpensive. The evaluation procedures would also correspond to the established mechanisms applied to multiple-choice tests, and a similar test duration could be expected.

A computer with monitor and mouse is required (alternatively, a tablet with touchscreen could also be used). The responses can be recorded by way of a passive medium, but the presentation of the test item requires a certain degree of interactive functionality (programming of objects which can be "swept aside").

## Schriftenreihe

## Berichte der Bundesanstalt für Straßenwesen

## Unterreihe „Mensch und Sicherheit“

# 2011

M 214: **Evaluation der Freiwilligen Fortbildungsseminare für Fahranfänger (FSF) – Wirksamkeitsuntersuchung**
Sindern, Rudinger € 15,50

M 215: **Praktische Fahrerlaubnisprüfung – Grundlagen und Optimierungsmöglichkeiten – Methodische Grundlagen und Möglichkeiten der Weiterentwicklung**
Sturzbecher, Bönninger, Rüdel et al. € 23,50

M 216: **Verkehrserziehungsprogramme in der Lehreraus-/Fortbildung und deren Umsetzung im Schulalltag – Am Beispiel der Moderatorenkurse „EVA“, „XpertTalks“, „sicherfahren“ und „RiSk“**
Neumann-Opitz, Bartz € 14,50

M 217: **Leistungen des Rettungsdienstes 2008/09 Analyse des Leistungsniveaus im Rettungsdienst für die Jahre 2008 und 2009**
Schmiedel, Behrendt € 16,50

M 218: **Sicherheitswirksamkeit des Begleiteten Fahrens ab 17. Summative Evaluation**
Schade, Heinzmann € 20,00

M 218b: **Summative Evaluation of Accompanied Driving from Age 17**
Schade, Heinzmann
Dieser Bericht liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

M 219: **Unterstützung der Fahrausbildung durch Lernsoftware**
Petzoldt, Weiß, Franke, Krems, Bannert € 15,50

# 2012

M 220: **Mobilitätsstudie Fahranfänger – Entwicklung der Fahrleistung und Autobenutzung am Anfang der Fahrkarriere**
Funk, Schneider, Zimmermann, Grüninger € 30,00

M 221: **Maßnahmen zur Erhöhung der Verkehrssicherheit von Kleintransportern**
Roth € 15,00

M 222: **Neue Aufgabenformate in der Theoretischen Fahrerlaubnisprüfung**
Malone, Biermann, Brünken, Buch € 15,00

M 223: **Evaluation der bundesweiten Verkehrssicherheitskampagne „Runter vom Gas!“**
Klimmt, Maurer € 15,00

M 224: **Entwicklung der Verkehrssicherheit und ihrer Rahmenbedingungen bis 2015/2020**
Maier, Ahrens, Aurich, Bartz, Schiller, Winkler, Wittwer € 17,00

M 225: **Ablenkung durch fahrfremde Tätigkeiten – Machbarkeitsstudie**
Huemer, Vollrath € 17,50

M 226: **Rehabilitationsverlauf verkehrsauffälliger Kraftfahrer**
Glitsch, Bornewasser, Dünkel € 14,00

M 227: **Entwicklung eines methodischen Rahmenkonzeptes für Verhaltensbeobachtung im fließenden Verkehr**
Hautzinger, Pfeiffer, Schmidt € 16,00

M 228: **Profile von Senioren mit Autounfällen (PROSA)**
Pottgießer, Kleinemas, Dohmes, Spiegel, Schädlich, Rudinger € 17,50

M 229: **Einflussfaktoren auf das Fahrverhalten und das Unfallrisiko junger Fahrerinnen und Fahrer**
Holte € 25,50

M 230: **Entwicklung, Verbreitung und Anwendung von Schulwegplänen**
Gerlach, Leven, Leven, Neumann, Jansen € 21,00

M 231: **Verkehrssicherheitsrelevante Leistungspotenziale, Defizite und Kompensationsmöglichkeiten älterer Kraftfahrer**
Poschadel, Falkenstein, Rinkenauer, Mendzheritskiy, Fimm, Worringer, Engin, Kleinemas, Rudinger € 19,00

M 232: **Kinderunfallatlas – Regionale Verteilung von Kinderverkehrsunfällen in Deutschland**
Neumann-Opitz, Bartz, Leipnitz € 18,00

# 2013

M 233: **8. ADAC/BASt-Symposium 2012 – Sicher fahren in Europa**
CD-ROM / kostenpflichtiger Download € 18,00

M 234: **Fahranfängervorbereitung im internationalen Vergleich**
Genschow, Sturzbecher, Willmes-Lenz € 23,00

M 235: **Ein Verfahren zur Messung der Fahrsicherheit im Realverkehr entwickelt am Begleiteten Fahren**
Glaser, Waschulewski, Glaser, Schmid € 15,00

M 236: **Unfallbeteiligung von Wohnmobilen 2000 bis 2010**
Pöppel-Decker, Langner
Dieser Bericht liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

M 237: **Schwer erreichbare Zielgruppen – Handlungsansätze für eine neue Verkehrssicherheitsarbeit in Deutschland**
Funk, Faßmann € 18,00

M 238: **Verkehrserziehung in Kindergärten und Grundschulen**
Funk, Hecht, Nebel, Stumpf € 24,50

M 239: **Das Fahrerlaubnisprüfungssystem und seine Entwicklungspotenziale – Innovationsbericht 2009/2010** € 16,00

M 240: **Alternative Antriebstechnologien – Marktdurchdringung und Konsequenzen – Berichtsjahr 2011 – Abschlussbericht**
Küter, Holdik, Pöppel-Decker, Ulitzsch
Dieser Bericht liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

M 241: **Intervention für punkteauffällige Fahrer – Konzeptgrundlagen des Fahreignungsseminars**
Glitsch, Bornewasser, Sturzbecher, Bredow, Kaltenbaek, Büttner € 25,50

M 242: **Zahlungsbereitschaft für Verkehrssicherheit – Vorstudie**
Bahamonde-Birke, Link, Kunert € 14,00

# 2014

M 243: **Optimierung der Praktischen Fahrerlaubnisprüfung**
Sturzbecher, Mörl, Kaltenbaek € 25,50

M 244: **Innovative Konzepte zur Begleitung von Fahranfängern durch E-Kommunikation**
Funk, Lang, Held, Hallmeier € 18,50

M 245: **Psychische Folgen von Verkehrsunfällen**
Auerbach € 20,00

M 246: **Prozessevaluation der Kampagnenfortsetzung 2011-2012 „Runter vom Gas!“**
Klimmt, Maurer, Baumann € 14,50

AKTUALISIERTE NEUAUFLAGE VON:
M 115: **Begutachtungsleitlinien zur Kraftfahreignung – gültig ab 1. Mai 2014**
Gräcmann, Albrecht € 17,50

M 247: **Psychologische Aspekte des Unfallrisikos für Motorradfahrerinnen und -fahrer**
von Below, Holte € 19,50

M 248: **Erkenntnisstand zu Verkehrssicherheitsmaßnahmen für ältere Verkehrsteilnehmer**
Falkenstein, Joiko, Poschadel € 15,00

M 249: **Wirkungsvolle Risikokommunikation für junge Fahrerinnen und Fahrer**
Holte, Klimmt, Baumann, Geber € 20,00

M 250: **Ausdehnung der Kostentragungspflicht des § 25a StVG auf den fließenden Verkehr**
Müller € 15,50

M 251: **Alkohol-Interlocks für alkohohlauffällige Kraftfahrer**
Hauser, Merz, Pauls, Schnabel, Aydeniz, Blume, Bogus, Nitzsche, Stengl-Herrmann, Klipp, Buchstaller, DeVol, Laub, Müller, Veltgens, Ziegler € 15,50

M 252 **Psychologische Aspekte des Einsatzes von Lang-Lkw**
Glaser, Glaser, Schmid, Waschulewski
Dieser Bericht liegt nur in digitaler Form vor, ist interaktiv und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

## 2015

M 253: **Simulatorstudien zur Ablenkungswirkung fahrfremder Tätigkeiten**
Schömig, Schoch, Neukum, Schumacher, Wandtner € 18,50

M 254: **Kompensationsstrategien von älteren Verkehrsteilnehmern nach einer VZR-Auffälligkeit**
Karthaus, Willemssen, Joiko, Falkenstein € 17,00

M 255: **Demenz und Verkehrssicherheit**
Fimm, Blankenheim, Poschadel € 17,00

M 256: **Verkehrsbezogene Eckdaten und verkehrssicherheitsrelevante Gesundheitsdaten älterer Verkehrsteilnehmer**
Rudinger, Haverkamp, Mehlis, Falkenstein, Hahn, Willemssen € 20,00

M 257: **Projektgruppe MPU-Reform**
Albrecht, Evers, Klipp, Schulze € 14,00

M 258: **Marktdurchdringung von Fahrzeugsicherheitssystemen**
Follmer, Geis, Gruschwitz, Hölscher, Raudszus, Zlocki € 14,00

M 259: **Alkoholkonsum und Verkehrsunfallgefahren bei Jugendlichen**
Hoppe, Tekaat € 16,50

M 260: **Leistungen des Rettungsdienstes 2012/13**
Schmiedel, Behrendt € 16,50

M 261: **Stand der Radfahrausbildung an Schulen und motorische Voraussetzungen bei Kindern**
Günther, Kraft € 18,50

M 262: **Qualität in Fahreignungsberatung und fahreignungsfördernden Maßnahmen**
Klipp, Bischof, Born, DeVol, Dreyer, Ehlert, Hofstätter, Kalwitzki, Schattschneider, Veltgens € 13,50

M 263: **Nachweis alkoholbedingter Leistungsveränderungen mit einer Fahrverhaltensprobe im Fahrsimulator der BASt**
Schumacher
Dieser Bericht liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

## 2016

M 264: **Verkehrssicherheit von Radfahrern – Analyse sicherheitsrelevanter Motive, Einstellungen und Verhaltensweisen**
von Below € 17,50

M 265: **Legalbewährung verkehrsauffälliger Kraftfahrer nach Neuerteilung der Fahrerlaubnis**
Kühne, Hundertmark € 15,00

M 266: **Die Wirkung von Verkehrssicherheitsbotschaften im Fahrsimulator – eine Machbarkeitsstudie**
Wandtner
Dieser Bericht liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

M 267: **Wahrnehmungspsychologische Analyse der Radfahraufgabe**
Platho, Paulenz, Kolrep € 16,50

M 268: **Revision zur optimierten Praktischen Fahrerlaubnisprüfung**
Sturzbecher, Luniak, Mörl € 20,50

M 269: **Ansätze zur Optimierung der Fahrschulausbildung in Deutschland**
Sturzbecher, Luniak, Mörl € 21,50

M 270: **Alternative Antriebstechnologien – Marktdurchdringung und Konsequenzen**
Schleh, Bierbach, Piasecki, Pöppel-Decker, Ulitzsch
Dieser Bericht liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

## 2017

M 271: **Evaluation der Kampagnenfortsetzung 2013/2014 „Runter vom Gas!"**
Klimmt, Geber, Maurer, Oschatz, Sülflow € 14,50

M 272: **Marktdurchdringung von Fahrzeugsicherheitssystemen 2015**
Gruschwitz, Hölscher, Raudszus, Zlocki € 15,00

M 273: **Verkehrswahrnehmung und Gefahrenvermeidung – Grundlagen und Umsetzungsmöglichkeiten in der Fahranfängervorbereitung**
TÜV | DEKRA arge tp 21 € 22,00

M 273b: **Traffic perception and hazard avoidance – Foundations and possibilities for implementation in novice driver preparation**
Bredow, Brünken, Dressler,Friedel,Genschow, Kaufmann,Malone, Mörl, Rüdel, Schubert, Sturzbecher, Teichert, Wagner, Weiße
Dieser Bericht ist die englische Fassung von M 273 und liegt nur in digitaler Form vor und kann unter http://bast.opus.hbz-nrw.de/ heruntergeladen werden.

Fordern Sie auch unser kostenloses Gesamtverzeichnis aller lieferbaren Titel an! Dieses sowie alle Titel der Schriftenreihe können Sie unter der folgenden Adresse bestellen:

**Fachverlag NW in der Carl Schünemann Verlag GmbH**
**Zweite Schlachtpforte 7 · 28195 Bremen**
**Tel. +(0)421/3 69 03-53 · Fax +(0)421/3 69 03-63**

Alternativ können Sie alle lieferbaren Titel auch auf unserer Website finden und bestellen.

**www.schuenemann-verlag.de**