The System of Driving Licence Testing and its Development Potential

Innovation Report on Optimisation of the
Driving Licence Test during the Report Period 2009/2010

TÜV | DEKRA arge tp 21 (Eds.)

The System of Driving Licence Testing and its Development Potential

Innovation Report on Optimisation of the

Driving Licence Test during the Report Period 2009/2010

Dresden

TÜV | DEKRA

CREDITS

Title:	**The System of Driving Licence Testing and its Development Potential** Innovation Report on Optimisation of the Theoretical Driving Licence Test during the Report Period 2009/2010
Editor:	TÜV | DEKRA Arbeitsgemeinschaft Technische Prüfstellen für den Kraftfahrzeugverkehr 21 (arge tp 21), Dresden

Address:	Wintergartenstr. 4 01307 Dresden, Germany Tel.: +49-(0)351-20789-0 Fax: +49-(0)351-20789-20 E-mail: sekretariat@argetp21.de www.argetp21.de

Authors and contributors:

Friedel, T.	TÜV | DEKRA arge tp 21
Genschow, J.	IPV Oberkrämer
Krampe, A.	IPV Oberkrämer
Rüdel, M.	TÜV | DEKRA arge tp 21
Schmidt, Dr., A.	DEKRA Automobil GmbH
Sturzbecher, Prof. Dr. D.	Universität Potsdam
Weiße, B.	TÜV | DEKRA arge tp 21

Citation suggestion: TÜV | DEKRA arge tp 21 (2011, Eds.). The System of Driving Licence Testing and its Development Potential - Innovation Report on Optimisation of the Driving Licence Test during the Report Period 2009/2010. Dresden: TÜV | DEKRA arge tp 21.

1st edition, 2011

Abstract

Innovation reports provide information on the course and results of research and development work conducted in connection with medium- and long-term further development of the system of driving licence testing during a given two-year period. With the aid of innovation reports, it is thus possible to assess the quality, progress and scientific basis for further development of the driving licence test.

The present innovation report describes the core aspects of research performed by the TÜV | DEKRA arge tp 21 working group with regard to the theoretical driving test over the report period 2009/2010. The defined priority tasks concerned (1) elaboration of a driving competence model, (2) evaluation and further development of the traditional question formats and test methodology, (3) research work on the use of computer-generated dynamic driving scenarios, and (4) elaboration of innovative test items for the testing of components of driving competence which have not been tested adequately to date in terms of action knowledge.

Re (1): Taking into account the content-based demand levels of driving behaviour (e.g. DONGES, 2009), alongside levels of driving competence acquisition (e.g. GRATTENTHALER, KRÜGER & SCHOCH, 2009), a "driving competence structure model" was developed to classify content-based components of driving competence and to structure the test items accordingly. At the same time, this permits elaboration of prototypical demand situations for the operationalisation of test content, as well as description of the areas of content and driving competence which can be covered by the different forms of testing.

Re (2): Following introduction of a PC-based theoretical driving test, the technical framework for test realisation (e.g. test program interface) was evaluated and optimisation potential suitable to reduce the susceptibility to manipulation was implemented (e.g. dissolution of the fixed pairs of basic and supplementary test sheets, randomised order of test questions and of the possible answers to individual questions). Building upon the results of the revision project, the continuous evaluation of test questions and parallel tests by the Institute for Prevention and Road Safety (IPV) demonstrated that, in principle, the vast majority of test questions in use fulfils the intended function of competence verification in accordance with various criteria relating to psychological testing.

Re (3): To improve the form of question presentation and the associated instruction formats, the TÜV | DEKRA arge tp 21 working group has developed the software solution "VICOM". With the aid of this software, the photographs used to date were replaced with computer-generated static images, which could hence be created and varied much more efficiently. Furthermore, it became possible to elaborate dynamic video sequences. The testing of questions with dynamic situation presentation suggests that the intended assessment of hazard perception competences could be better achieved by the new instruction format with no solution hints remaining visible in the final image (FRIEDEL, WEIßE & RÜDEL, 2010).

Re (4): Development potential for the theoretical driving test is identified especially with regard to the assessment of action competences in the fields of traffic perception and hazard avoidance. At the time of testing, the candidate does not usually possess the driving experience which can be considered prerequisite for acquisition, and thus these safety-relevant competences cannot be addressed in the traditional "knowledge test". The possibilities for assessment within the traditional practical test are likewise limited, as the demand situations in real traffic cannot be controlled at will by the examiner, and hazardous situations also cannot be brought about deliberately for safety reasons. It appears necessary, therefore, to develop an innovative form of testing for the German driving licence test which would permit realistic computer simulations of traffic and, in particular, hazard situations, and could thus be used to operationalise the aforementioned competence components. Corresponding hazard perception tests are already in use in the systems of novice driver preparation in a number of European and overseas countries.

Further realisation of the optimisation potential for driving licence testing in Germany requires that its further development be placed in the context of the overall system of novice driver preparation. To raise the efficiency of novice driver preparation, the quality assurance and further development measures must be focussed not merely on input specifications such as teaching plans and examination guidelines, but instead also on output specifications such as the competence level to be achieved by the novice driver. The training standards to be defined must describe the minimum levels of driving competence which a novice driver should have achieved at the transitions between the individual phases of novice driver preparation in a sufficiently specific manner to permit translation into test items, and thus evaluation within the framework of driving licence tests.

Contents

Bernd Weiße, Dietmar Sturzbecher & Mathias Rüdel

1 Innovation reports as a means for further development of the theoretical driving test

1.1 Historical background to further development of the theoretical driving test over the past decade

Overall road accident figures at the end of the 1990s indicated a continuous decline in the numbers of persons injured or killed in road traffic. At the same time, however, it was shown that the risk of road traffic injury or death for novice drivers was still several times greater than for experienced drivers. Against this background, it can be asked whether the German system makes optimum use of the possibilities to prepare novice drivers for independent participation in motorised road traffic. The Federal Highway Research Institute (Bundesanstalt für Straßenwesen, BASt) reacted to this open question by initiating an investigation of the potential for possible optimisation of single aspects of novice driver preparation[1] with regard to their contribution to a reduced accident risk for novice drivers. The federal ministry responsible for road traffic later declared reductions in the number of accidents involving young novice drivers to be a key objective of its road safety improvement programme of February 2001 (WILLMES-LENZ, 2002).

The driving licence test is an especially important element within the overall system of novice driver preparation: On the one hand, it serves 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 contents, assessment criteria and test results provide important orientation for the organisation of driving school training and the individual learning processes of the novice drivers ("control function"). Given this special significance, the driving licence test formed the nucleus of work on optimisation of the overall system of novice driver preparation from the very beginning. Already in 1997, for example, the BASt announced the project "Optimisation of driving licence testing", which was finally processed by the Technical Examination Centres for Motor Vehicle Traffic (Technische Prüfstellen für den Kraftfahrzeugverkehr), the bodies mandated to conduct and further develop driving licence tests in Germany[2]. It soon became clear, however, that an adequate treatment of the project topic was not feasible with the allocated means and personnel resources. Consequently, the participants agreed to concentrate initially on the evidently greater optimisation potential in the theoretical driving test (hereafter also TDT) and to address possible optimisation of the practical driving test (hereafter also PDT) in a separate project; the latter project was then financed and undertaken independently by the Technical Examination Centres between 2005 and 2008 (STURZBECHER, BÖNNINGER & RÜDEL, 2008, 2010). The potential for TDT optimisation identified within the framework of the BASt project "Optimisation of driving licence testing" is described in the corresponding project report (BÖNNINGER & STURZBECHER, 2005). It is seen above all in the provision of a stronger scientific basis for evaluation and further development of the TDT, and in exploitation of the benefits of computer-assisted test methods.

Realisation of this optimisation potential subsequently began with a pilot project and a revision project handled by the Technical Examination Centres. These projects served to establish the technical and scientific prerequisites for the introduction of a computer-assisted TDT and the associated scientific evaluation. The work culminated in a policy decision of the Federal/Regional Expert Committee "Driver Licensing and Driving Instructor Legislation" (Bund-Länder-Fachausschusses "Fahrerlaubnisrecht/Fahrlehrerrecht", BLFA-FE/FL) on optimisation of the TDT on 12th October 2007. The necessary amendment of the Schedule of Fees for Measures Concerning Road Traffic (Gebührenordnung für Maßnahmen im Straßenverkehr, GebOSt) was passed by the Bundesrat at its 840th session on 20th December 2007, whereby the scientific foundation and continuous further development of driver testing were shifted into the focus. Both were ensured through the introduction of new methods and procedures, as described in the ensuing "System Manual on Driver Licensing (Theory Test)" ("Handbuch zum Fahrerlaubnisprüfungssystem (Theorie)", TÜV | DEKRA arge tp 21, 2008). This manual serves all organisations involved in the system of driving licence testing as a common basis for the realisation and further

[1] "Novice driver preparation" is here understood to mean the entirety of all conditions and measures which are laid down in legislation or, beyond that, provided and used specifically in a particular cultural context to permit the learning of independent, safe and responsible driving of a motor vehicle in public road traffic and demonstration of the necessary knowledge and ability (see Chapter 1.3).

[2] The historical development of the driving licence test and the interactions between the individual involved organisations are described in the report "The History of the Driving Test in Germany" (BÖNNINGER, KAMMLER & STURZBECHER, 2009).

development of the TDT and came into effect after government acknowledgement at federal and state level on 6th November 2008.

1.2 Function and purpose of the innovation reports

Regular reporting on the course of evaluation and further development processes is an essential element of the optimised driving licence testing system and corresponding reformation of the TDT. In this context, innovation reports are to be viewed as a core form of documentation with strategic importance.[3] Issued every two years, they document all significant research and development processes in connection with the 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)”, the reports 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 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 with improved test procedures, and 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 working group “TÜV | DEKRA arge tp 21”[4]. 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 further development of the driving licence test. 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 2009 to 2010

The starting positions for the research work described in the present report can be derived from

- the fundamental function of the TDT in the overall system of novice driver preparation,
- the development status of novice driver preparation in Germany during the report period in general, and the status of the TDT in particular, and
- the potential for optimisation of the TDT identified to date, alongside assessment of the possibilities and priorities for effective realisation of this potential.

These starting positions are explained in more detail in the following sections.

[3] 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.

[4] The TÜV | DEKRA Working Group “Technical Examination Centres in the 21st Century” (“TÜV | DEKRA Arbeitsgemeinschaft der Technischen Prüfstellen im 21. Jahrhundert”, here abbreviated for the sake of better legibility to “TÜV | DEKRA arge tp 21”) was founded in 1999 as a joint venture of the operators of the Technical Examination Centres in Germany, namely TÜV Rheinland Kraftfahrt GmbH, TÜV SÜD Auto Service GmbH, TÜV NORD Mobilität GmbH & Co. KG and DEKRA Automobil GmbH.

Function of the TDT in the overall system of novice driver preparation

The long since internationally recognised term “novice driver preparation” has only been current in the terminology of the German research community for a few years. Genschow, Sturzbecher & Willmes-Lenz (in press) gave precision to the definition within the framework of the BASt project “Novice Driver Preparation in Europe” and pointed out that the term can be considered under both socialisation and institutional aspects.

With regard to the former, it must be noted that the preparation of novice drivers for their independent participation in motorised road traffic is accomplished over the course of a long-term (traffic) socialisation process[5] which already begins in early childhood and in which driving licence acquisition represents merely one period. Such a broad understanding of the concept certainly appears promising from the perspective of traffic pedagogy; for studies concerning driver licensing, and particularly driving licence testing, however, an institutional approach is to be recommended, as it also takes into account the interdependencies between legislative regulations, safety-relevant traffic and driving competences, and the demands of examination didactics typical in this field. At the same time, due consideration is given to the close interactions between state authorities, the mandated Technical Examination Centres, commercial driving schools and all other persons and organisations involved in novice driver preparation.

If “novice driver preparation” is perceived as an institution within the aforementioned socialisation process, then it provides regulative, normative and cultural-cognitive structures (the three so-called “pillars” of an institution)[6] to support an individual's development. These structures contribute to social stability (in the present context to road safety), because they establish binding obligations and rights to govern the actions of the participants in road traffic. Legislative instruments (e.g. the Road Traffic Act or Driving Licence Regulations) defining the framework for driving competence acquisition (e.g. stipulations on test contents, minimum ages or special conditions) constitute the regulative structure of the first pillar. Values such as environmentally aware driving and the assumption of responsibility towards “weaker” road users, as well as overarching aims such as an improvement in road safety, represent the normative structure and the second pillar. The contents and forms of driving and traffic competence acquisition and the organisations involved in this acquisition process (e.g. schools, driving schools, Technical Examination Centres or the police), finally, can be assigned to the sphere of cultural-cognitive structures and thus to the third pillar. In this sense, novice driver preparation can be viewed as an institution which serves to promote and formalise (driving) socialisation; this accentuation of the concept leads to a certain strengthening of the focus on driving licence acquisition, without losing sight of the adjacent aspects of competence acquisition. This includes the diverse teaching and learning forms by which driving competence is acquired, as well as the different forms of testing serving to furnish proof of such competence.

On the basis of studies of teaching and learning theory and an analysis of the systems of novice driver preparation in 44 countries, Genschow, Sturzbecher and Willmes-Lenz (in press) distinguish four forms of testing alongside a series of prescribed or otherwise typical teaching and learning forms (e.g. theory classes, independent theory learning, practical driving instruction, accompanied driving, advanced training courses and solo driving under protective regulations), namely the forms “knowledge test” (or “theory test”), “traffic perception test” (or “hazard perception test”), “learner assessment” and “(practical) driving test”. In the present context, knowledge tests serve specifically to demonstrate an adequate scope of driving- and traffic-related knowledge and have to date usually limited themselves to the verification of knowledge formulated in test questions pertaining to rules and facts (declarative or factual knowledge), whereas action-related knowledge application (procedural or action knowledge) has often played no particular role.

The questions in knowledge tests generally follow standardised formats (e.g. multiple-choice questions, true/false questions, gap-fill questions). With the aid of visual media and correspondingly innovative formats, however, it would be possible to extend the current spectrum and to assess the application of knowledge (i.e. action knowledge) to a much greater extent. Computers with multimedia capabilities hold particular potential, as they are able to depict driving and traffic situations in a near-realistic manner and

[5] Hurrelmann (1999, p. 481) uses the term ‘socialisation’ to describe “… the process of personality formation characterised by the mutual dependence and permanent interactions of a socially conveyed social and material environment on the one hand and the biophysical structure of the organism on the other.”

[6] According to Scott (1995), the regulative pillar stands for those aspects of institutions which limit and regulate our actions. The focus is here placed on the stipulation of rules and on the monitoring and sanctioning of behaviour. The normative pillar represents the evaluative and obliging dimension of institutions; this refers not only to values, but also to objectives and the conventions to be observed for their attainment. The cultural-cognitive pillar, finally, reflects the manner of the perception and acquisition of “reality” in a society; this is influenced by cultural factors.

can furthermore record and evaluate non-verbal responses from the novice driver. The latter approaches can be classified more adequately as “traffic perception tests”. Traffic perception and hazard recognition are here the test categories placed in the foreground. The predominant test method is to demand a correct reaction or the correct “driving decision” in the displayed scenario; at the same time, non-verbal response is also measured (e.g. the reaction time before a computer input). The computer is for traffic perception tests the essential medium for task presentation and the recording of reaction times.

In the international field, it can be seen that progressive systems of novice driver preparation increasingly incorporate hazard perception tests, either as an independent form of testing alongside the traditional knowledge test or as innovative, supplementary formats. The traditional knowledge test is thus experiencing further development and differentiation on the basis of the availability of new media; in this way, its weaknesses in the assessment of subject matter relating to hazard avoidance are increasingly overcome, thanks to innovative forms of testing in which dynamic driving scenarios provide visual instruction to illustrate the task in hand (GENSCHOW, STURZBECHER & WILLMES-LENZ, in press).

The acquisition of action competence in general, and driving competence in particular, begins with the systematic development of flexible, connectable and transferable knowledge of the subject matter underlying the particular action, which in the present case means the circumstances of motorised road traffic (first step). On this basis, it is then necessary to acquire the ability to apply the knowledge concerned effectively and in a manner appropriate to an action situation, i.e. in the contexts of diverse traffic situations (second step). The final outcome of this process is the accumulation of a differentiated repertoire of problem- and situation-related action patterns, from which the immediately appropriate (driving) behaviour can be called up (third step). In the light of this three-step process of competence or expertise acquisition (cf. ANDERSON, 2001; GREENO, COLLINS & RESNICK, 1996; GRUBER & MANDL, 1996), traditional knowledge tests, which address above all declarative knowledge and thus equate to the first step, can be placed at a relatively early stage of the process of novice driver preparation, and indeed must be arranged so, if they are to provide meaningful motivation and support to knowledge acquisition within this process. The greater operationalisation of procedural knowledge in further developed knowledge tests or new traffic perception tests, on the other hand, suggests placement of such forms of testing at a later point in the course of novice driver preparation due to the correspondingly longer learning processes (second and third steps).

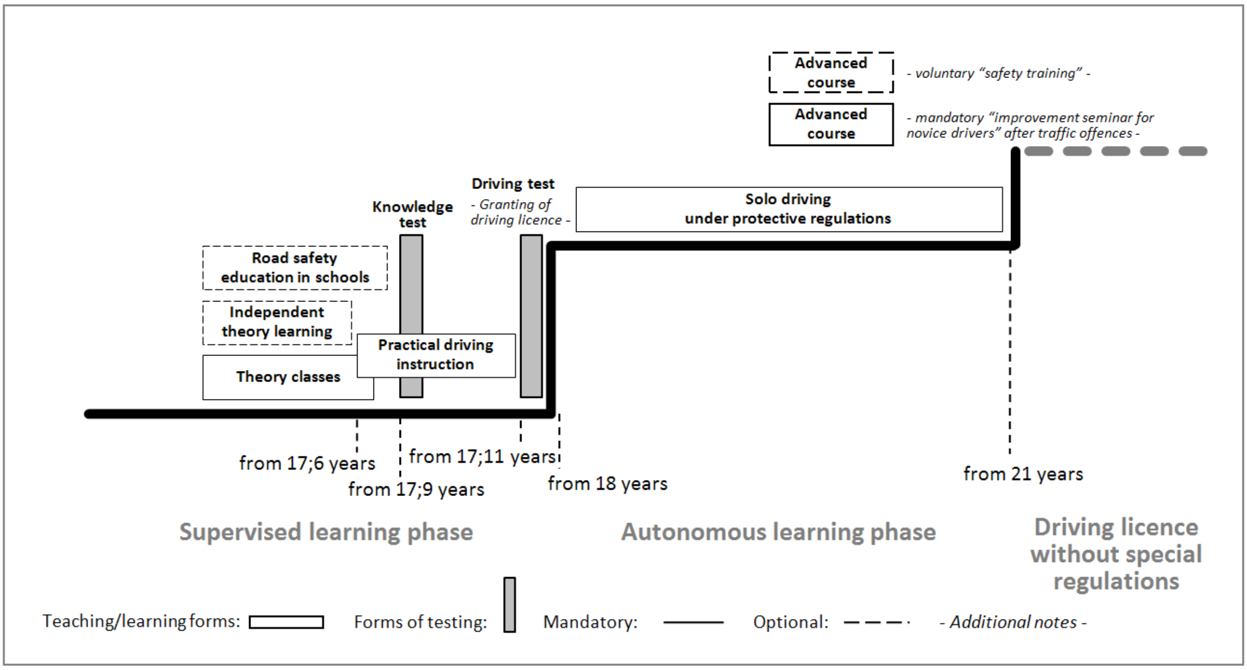

Fig. 1 System of novice driver preparation in Germany – Model with solely driving school training (GENSCHOW, STURZBECHER & WILLMES-LENZ, in press)

Development status of novice driver preparation in Germany at the beginning of the report period in general and the status of the TDT in particular

During the period covered by the report, novice drivers were able to choose between driver training solely in a driving school and the model "Accompanied driving from age 17" ("AD17") ("Begleitetes Fahren ab 17", "BF17"). In the former case (see Fig. 1), the novice drivers attend the prescribed theory classes and complete the prescribed practical instruction with a professional driving instructor during the supervised learning phase, commencing at the earliest at an age of 17 years and 6 months. They must first complete a knowledge test ("theoretical driving test") and finally a test drive ("practical driving test"). Passing the practical driving test founds an entitlement to drive solo (at the earliest from an age of 18 years).[7]

In case of participation in the AD17 model (see Fig. 2), the novice drivers must likewise attend formal driver training with theory classes and practical instruction in a driving school; subsequently, they must also complete a theoretical and practical driving test. The earliest possible dates for the commencement of driving school instruction and for the associated tests, however, are one year earlier for participants in the AD17 model compared to the model of solely driving school training. After passing the driving test, the AD17 participant is initially only permitted to drive a motor vehicle when accompanied by an experienced licence holder (at the earliest from the age of 17 years, and then up to the age of 18 years).

In both models, the novice driver is issued a "probationary driving licence", which is subject to a number of protective regulations during the first two years[8] after passing the driving test. In the case of the AD17 model, the novice driver is thus already deemed the legally responsible driver of a vehicle from the beginning of the period of accompanied driving practice, despite the stipulation of an experience accompanying passenger.

Already at the beginning of the report period in 2009, the theoretical driving test was being conducted increasingly as a computer-based test in Germany. The test for a driving licence for vehicle class B comprises a total of 30 test items (multiple-choice questions where at least one of the offered answers is correct, and gap-fill questions requiring a numerical value to be entered). The questions are weighted with

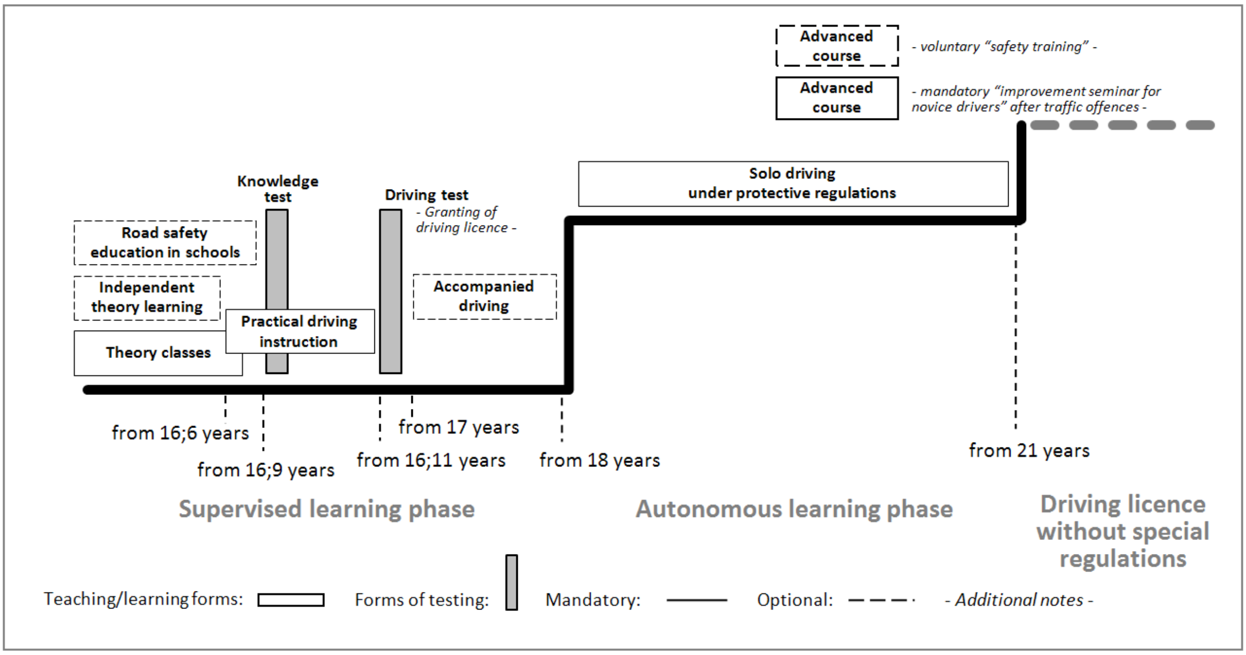

Fig. 2 System of novice driver preparation in Germany – Model of "Accompanied driving from age 17 (AD17)" (GENSCHOW, STURZBECHER & WILLMES-LENZ, in press)

[7] The diagram shows only the (legally) stipulated minimum ages for driving licence acquisition and the associated prescribed or typical arrangement of the elements of preparation. The lengths of the phases depicted in the diagram are not proportional to the durations of driving competence and driving licence acquisition.

[8] The "zero alcohol for novice drivers" rule which came into effect on 1st August 2007 applies during the two-year probationary period and thereafter, where appropriate, until the driver reaches the age of 21 years.

different score penalties for incorrect answers; the test is considered passed if a total penalty of 10 points is not exceeded, provided no more than one question weighted with a score penalty of 5 points is answered incorrectly. The instructions for certain questions are illustrated with drawings or photographs, with traffic situations being depicted from the perspective of the driver. The test questions cover the subject areas “Hazards”, “Behaviour in road traffic”, “Right of way”, “Traffic signs”, “Environment protection”, “Regulations on the use of motor vehicles”, “Technical aspects of a motor vehicle” and “Qualification and fitness to drive”. Each of these subject areas is broken down further into subcategories. The result of the test is announced in person by the driving test examiner at the end of the test. The novice driver is informed as to the questions which were answered correctly and any subject areas in which learning deficits persist. If the test is not passed, it can be repeated at the earliest after two weeks. The pass rate in the theoretical driving test lies at around 75 per cent.

In summary, it can be stated that the theory test in the German system of novice driver preparation is conducted at a relatively early stage after a short period of basic theoretical instruction and independent theory learning – although nothing is known about the scope and individual contents of the independent theory learning. The test contents thus concentrate on the field of declarative (factual) knowledge. The aforementioned innovative test item formats, which enable the use of dynamic driving scenarios for content presentation and new possibilities for the testing of traffic and hazard perception, still played no role at the beginning of the report period. The TDT was at this time still to be viewed as a traditional knowledge test; there were no perception tests or at least corresponding innovative item formats. This circumstance is in itself an indication of potential for optimisation and further development of the TDT. The significance of these opportunities must not be underestimated, because the TDT is positioned relatively closely to the transition from the protected phase of practical training under the supervision of a driving instructor to the phase of solo (or initially accompanied) driving. From the perspective of driving safety, this transition is the crucial reference point for safety-related considerations, as it is during the first period of solo driving that novice drivers face by far the greatest risk of accident involvement of their whole driving career (“initial peak of endangerment”). In terms of its control and selection function, the TDT is alongside the PDT decisive in determining whether or not a novice driver has acquired sufficient driving competence to master the risks of solo driving. To better satisfy this function, it would be desirable to address a greater proportion of procedural knowledge in the test contents, in particular with regard to the possibilities of hazard perception and avoidance.

Potential for optimisation of the TDT present at the beginning of the report period and specification of the priorities for realisation

The extent to which the TDT satisfies the demands placed on a modern knowledge test was investigated immediately prior to the report period within the framework of a revision project of the Technical Examination Centres. The results of the evaluation were summarised in a revision report (STURZBECHER, KASPER, BÖNNINGER & RÜDEL, 2008), which also serves as the basis for various further research topics to be explained in detail in the following. The revision project results at the same time furnished proof of the effectiveness and validity of the evaluation model outlined and proposed by Bönninger and Sturzbecher (2005). With this backing, the Technical Examination Centres and regulatory authorities continued their further elaboration and development of the evaluation processes and defined corresponding basic principles in the “System Manual on Driver Licensing (Theory Test)”.

The revision report also contains numerous recommendations which are no less valid today with regard to the content and methodical optimisation of the TDT. Test questions relating to the subject areas “hazard avoidance” and “affective-emotional behaviour in road traffic”, especially, were found to be in considerable need of revision. These questions generally display a high probability of correct answering, which, in many cases, could apparently be attributed to particularities in the question formulation (e.g. formal properties of the question design which function as “solution hints”). It must be noted, however, that the weaknesses of the multiple-choice format used in the TDT with regard to the operationalisation of questions assessing hazard recognition and hazard management have been known for a long time and have already been the subject of optimisation efforts (HAMPEL, JANITSCHKE & SCHAFFRAN, 1977, HAMPEL, 1977); on the other hand, it is only since the advent of modern computer systems that these weaknesses can be countered effectively.

To alleviate the deficits revealed by the evaluations performed within the framework of the revision project, the development and evaluation group (Entwicklungs- und Evaluierungsgruppe, EEG) which is lead-managed by the working group TÜV | DEKRA arge tp 21 and brings together representatives not only of the Technical Examination Centres, but also of the competent federal and regional authorities, the Association of Technical Inspection Agencies (Verband der Technischen Überwachungsvereine, VdTÜV), the Federal Armed Forces (Bundeswehr) and the German Federation of Driving Instructor Associations

(Bundesvereinigung der Fahrlehrerverbände, BVF), has identified optimisation of the content and methodical design of conspicuous test items as a priority task. The more or less obvious solution hints contained in the answer options, in particular, are to be eliminated in the two aforementioned subject areas, and the overall challenge level of these items is to be raised. The increased use of illustrations or dynamic (virtual-reality) driving scenarios in instructions and replacement of the traditional multiple-choice format with innovative, audio-visual test items were also seen as concrete possibilities for optimisation. These possibilities originated from a resolution on the nationwide introduction of a computer-assisted theoretical driving test in 2007 (BUNDESRAT, 2007) and were now to be investigated by the TÜV | DEKRA arge tp 21 working group during the report period.

It was finally decided by the Technical Examination Centres that the research and development potential of the TÜV | DEKRA arge tp 21 working group was to be concentrated on a question already posed in the report on the BASt project "Optimisation of driving licence testing" (BÖNNINGER & STURZBECHER, 2005), namely how the possibilities of the computer as a test medium could be used – in addition to the aforementioned traffic perception tests – to assess further aspects of driving and traffic competence. The work to exploit such optimisation potential, however, quickly led to further questions which go beyond the realms of the TDT and influence the whole process of novice driver preparation; these questions are to be mentioned only briefly at this point:

- Which aspects of driving competence can be assessed on a PC or driving simulator, and where is this (only) possible by way of a test drive in an actual vehicle?
- Which teaching and learning forms are possibly prerequisites for the use of new forms of testing?
- Where should (new) forms of testing be positioned within the processes of driving competence acquisition and novice driver preparation in order to achieve the greatest possible effects with regard to improved driving safety?

The consideration of such questions automatically expands the perspective of the working group TÜV | DEKRA arge tp 21 from the TDT alone to the overall system of novice driver preparation. Coordinated further development of all elements of novice driver preparation, however, requires correspondingly founded assumptions regarding the structure and course of a process of driving competence acquisition. These assumptions are generally formulated in the form of competence models. The importance attached to such competence models has increased significantly over the past few years, in line with a paradigm shift from input to output control in the steering of education systems, above all with regard to schools (KLIEME et al., 2007). The essence of this shift is that quality assurance and further development measures are focussed less on input specifications such as teaching plans and examination guidelines, and instead on output specifications such as the competence level to be achieved by the target group. The stronger output orientation not only promises greater efficiency for the education systems, but also corresponds to a changing perception of the state, which limits itself to control functions and relieves the protagonists by defining education quality not in the form of detailed guidelines and regulations, but rather by way of objectives and the monitoring of their attainment.

The optimisation strategies indicated here can also be applied to the system of novice driver preparation, not least because its fundamental characteristics – institutional basis, competence development as the central objective, combination of formal and informal learning offers, independent central examination board, state monitoring, etc. – constitute a dedicated education system and establish parallels to the school system. A number of benefits could be expected for the system of novice driver preparation from increased output orientation, for example

- strengthening of the content-related interactions in the overall system and between its elements,
- expansion of the possibilities to test driving competence, and
- improved professionality with regard to psychological testing and the methodical assessment of partial competences.

A basis for such stronger output orientation could be provided through the elaboration and introduction of corresponding training standards (KLIEME et al., 2007), which in the context of novice driver preparation must meet the following requirements:

- They must formulate binding requirements for the overall system of novice driver preparation, thereby connecting and coordinating the training and test processes.
- They must provide a precise and clear definition of the essential objectives of novice driver preparation in the form of the desired learning achievements to be demonstrated by novice drivers. This definition lends concrete substance to the "training mandate" to be fulfilled by the system of novice driver preparation.

– They must describe the minimum levels of driving competence which a novice driver should have achieved at the transitions between the individual phases of novice driver preparation. These graduated levels must be described specifically enough to permit translation into test items, and thus evaluation within the framework of driving tests.

To be able to elaborate such training standards for novice driver preparation, it is first necessary to develop a driving competence model, in which training and test contents are defined for each element or stage of novice driver preparation (structure model) and the levels to be achieved for each component of driving competence are described (level or process model). In the ideal case, the ensuing model of driving competence should be empirically validated and standardised. Compliance with this aim represents a major challenge, and is also yet to be solved for the school education system, despite very intensive research efforts.[9] Nevertheless, it is a challenge which should still be tackled in connection with further development of the system of driving licence testing.

1.4 Objectives of the present innovation report

Starting out from the positions sketched above, the TÜV | DEKRA arge tp 21 working group began definition of the priority aspects for research aimed at further development of the TDT in 2009. These issues were treated during the report period in cooperation with relevant scientific institutes and are described in the present innovation report in respect of the associated development processes, the results obtained and the necessary further research and development strategies:

1. The first key task was research and idea development regarding the elaboration of a driving competence model. The results of this theoretical work are summarised in Chapter 2.
2. Evaluation and further development of the traditional question formats and of the test methodology formed a second focus for work over the past two years. This task was derived directly from the specifications and results of continuous evaluation of the TDT and relates to the optimisation demands to be treated in the short and medium terms. The studies and measures realised in this connection are to be outlined with the results in Chapter 3, enabling the identification of follow-up work topics for the future. This refers above all to the optimisation of competence assessment within the framework of the existing multiple-choice question format, the utilisation of previously unexploited evaluation strategies to optimise question instructions and question grouping, and measures to reduce the susceptibility to manipulation.
3. The described weaknesses of the multiple-choice question format – with text instructions supplemented, where appropriate, by static illustrations – must be overcome in the medium term, above all with regard to the operationalisation of content from the subject area of hazard avoidance. This third key task is addressed in Chapter 4 of the present innovation report. The intention is here to describe fundamental research work on the use of a diversity of computer-generated dynamic driving scenarios serving as a prerequisite both for new instruction formats in connection with conventional multiple-choice questions and for innovative test items to assess traffic perception. The policy decision reached by the Federal/Regional Expert Committee “Driver Licensing and Driving Instructor Legislation” (BLFA-FE/FL) in March 2011, namely gradual introduction of test items with dynamic situation presentation from 2012, in addition to the existing question forms, emphasises the need for short-term elaboration and testing of such items.
4. The fourth priority task area, finally, concerned the necessity to elaborate innovative test items for the testing of components of driving competence which have not been tested to date. This relates above all to content with explicit reference to traffic perception in general and to hazard recognition and avoidance in particular. The first research steps in pursuit of this long-term development objective were undertaken during the report period. The corresponding deliberations are presented in Chapter 5 and discussed with regard to their significance for further development of the TDT (e.g. for the elaboration of new forms of testing, such as traffic perception tests, and for their placement within the system of novice driver preparation).

[9] The school education system has been facing up to this challenge by way of a priority programme of the German Research Foundation (DFG) entitled “Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes” since 2007.

Dietmar Sturzbecher & Bernd Weiße

2 Possibilities for the modelling and measurement of driving competence

2.1 Types and functions of driving competence models

As a prerequisite for targeted further development of the driving licence test within the framework of novice driver preparation, it is first necessary to answer a number of important questions:

- "What knowledge and ability is necessary to be able to drive a motor vehicle in road traffic?"
- "Which minimum level of knowledge and ability must be present to reasonably exclude risks for the driver and other traffic participants?"
- "Which aspects of this knowledge and ability can be assessed reliably and validly with different forms of testing?"

The first question demands initially an overarching theoretical description of those content components of knowledge and ability which are required for participation in motorised road traffic. This leads on to a concept and structural model of "driving competence". On this basis, it is then possible to develop competence level and competence acquisition models in answer to the second question. The third question, finally, addresses the method options and the limitations of different forms of testing as means to assess the attained level of particular components of competence. These correlations are to be explained in closer detail in this chapter.

Concept of driving competence

Definition of the concept of "driving competence" can only follow from a more general theoretical understanding of competence. The theoretical roots of this concept are to be found in action theory in the model of "vocational action competence", which has served to describe the demands of the vocational world since the early 1970s (ROTH, 1971). Subsequently, the concept of competence was used in various educational contexts and with different meanings. To sharpen perceptions of the concept of competence and to enhance its potential benefits for further development of the education system, Weinert's influential OECD report of 1999 for the first time presented a systematic overview of the competence concepts in use (in the social sciences) and – after weighing up different theoretical standpoints and empirical findings from cognitive and developmental psychology – recommended that, in the context of education, competences be viewed as functionally determined cognitive performance dispositions which are related to particular classes of situations and demands and can be described as knowledge, skills, strategies, routines and domain-specific abilities.

On this basis, Weinert (2001) later argues furthermore that the most workable definition of competence is that which was developed in the field of expertise research. Expertise research stresses the importance of subject-specific knowledge and practice-related experience for the acquisition of competence in selected subject areas, which are designated "domains". Competencies are in this interpretation the dispositions which enable a person to recognise and solve certain categories of problems successfully, in other words to master specific demand situations of a particular type (e.g. the safe driving of a motor vehicle in traffic). If novice driver preparation is understood as a practice-oriented and experience-driven socialisation process and as an educational institution (see above), then the sketched starting positions can also be deemed valid here. Consequently, in conformity with Weinert (2001, p. 27), "driving competence" can be defined as those cognitive abilities and skills which are available to or can be learned by an individual as a means to solve certain problems, alongside the motivational, volitional and social readiness und ability to realise the problem solutions successfully and responsibly in variable traffic situations (STURZBECHER, 2008 and 2010).

The above definition of driving competence – like the general concept of competence suggested by Weinert – thus refers to demand situations in a particular life context. The life context in the present case is motorised road traffic, which according to Sturzbecher (2008, 2010) constitutes a poorly defined or "lifeworld" domain with its changing conditions (e.g. weather conditions, traffic density). Such lifeworld domains are characterised by a high level of complexity and dynamism, i.e. constantly varying demands subject to external influencing factors. There are thus no rules or principles with equal validity for the response to all demand situations; it is rather that a specific problem solution strategy must be generated for each individual demand situation. Klieme et al. (2007) derive two requirements for the description or diagnosis of competences from their context specificity and situation reference. These requirements must also be taken into account when modelling and assessing driving competence (and thus in the further development of driving licence testing):

- Each operationalisation of a competence must refer to specific classes of demand situations.
- The scope of demand situations must mirror a broad performance spectrum (facet structure).

Driving competence acquisition

Transferable ("intelligent") knowledge suitable for flexible application, also in new situations, is to be viewed as the basis for all competence (BAUMERT, 1993). The acquisition of competence thus begins with the building of systematic intelligent knowledge in a particular domain; it is best promoted by a mix of systematic and situated learning, i.e. learning in real-life situations (WEINERT, 1998). Two different forms of knowledge are distinguished: Declarative or factual knowledge and procedural or action knowledge. These two forms of knowledge, however, are indivisible with regard to their acquisition and function. Procedural knowledge, for example, builds upon a foundation of declarative knowledge: A skill such as gear changing is initially performed as a knowledge-based action, before becoming automated or proceduralised.

The successful processing of complex tasks such as the driving of a motor vehicle demands an integrated utilisation of both declarative and procedural knowledge aspects in combination with the remaining elements of competence: "Qualification to perform a task means not only possessing the necessary declarative knowledge, but also prior acquisition of a cognitive system bringing together consciously accessible information, highly automated skills, intelligent strategies for knowledge application, a feel for the scope and quality of the available knowledge, positively realistic self-evaluation, and finally the action and learning motivation inherent to the individual's competence" (WEINERT, 1998, p. 111). This quote also provides a concise and pertinent description of the requirements to be met during the acquisition of driving competence.

Driving competence models

Taking the described basic positions in competence theory, it follows that driving competence models refer to the demand situations of motorised road traffic and should reflect these situations as comprehensively as possible. Assuming that driving competence models are to serve as a basis for training standards in the overall system of novice driver preparation, as explained in Chapter 1, they must furthermore fulfil the following functions (KLIEME et al., 2007, p. 74):

- Description of the content structure of those demands which novice drivers are expected to satisfy (components of driving competence), and
- Development of scientifically founded notions of the possible graduations of driving competence, i.e. the competence levels which can be determined in the case of novice drivers (stages of driving competence).

In accordance with these two functions, Klieme and Leutner (2006) distinguish competence structure models, which deal with the question of "which and how many distinct dimensions of competence can be differentiated in a specific field", and competence level models, which refer to the "concrete situative demands which a person is able to master given a particular scope of competence" (p. 6). Although these models relate to different aspects of the competence construct (content structure and attained levels), they are not mutually exclusive and are ideally even complementary (KOEPPEN, HARTIG, KLIEME & LEUTNER, 2008). Consequently, it would be desirable for an ideal model of driving competence for the context of novice driver preparation to describe both the various content-based dimensions of driving competence (partial competences) and the possible levels of attainment of such partial competences among novice drivers.

Both the structure models and the level models of driving competence can be differentiated further. In the case of structure models, Schecker and Parchmann (2006) distinguish between (1) normative competence models, which indicate the (cognitive) prerequisites to be met by a learner to be able to solve tasks and problems in a particular subject or demand field, and (2) descriptive competence models, which describe "typical" patterns of (cognitive) prerequisites with which it is possible to map the behaviour of a learner when solving such subject- and demand-specific tasks and problems. In respect of level models, it must be taken into account whether the described levels of driving competence represent merely the possible manifestations of driving competence or also the steps leading to their acquisition; in the latter case, it is customary to speak of “competence acquisition models”[10]. Independently of the model type, theoretically developed competence models can initially only be considered hypotheses and must subsequently be verified empirically (KLIEME, 2004).

Already from this brief overview of the different types of competence model and their specific functions, it is clear that there can be no complete or general model of driving competence which fulfils all functions in ideal manner. This is also true because models always serve a certain descriptive purpose and thus focus on very specific aspects of a phenomenon in reality. For the description of driving competence, it would naturally be desirable to develop all the above model types with mutual reference as far as possible (i.e. to elaborate acquisition models for the various components of driving competence, for example); this, however, is a time-consuming and very complex process.

The subject to be handled by the TÜV | DEKRA arge tp 21 working group is further development of the driving licence test. The aim of the recently begun work on a driving competence model is thus to systematise the contents and methods of the various current and future forms of testing in a normative model. This model is to be built initially on the relatively general level of structure and acquisition aspects, and then expanded into further detail. This development focus stemmed above all from the necessity to define short, medium and long-term objectives for the work on optimisation of the driving licence test. The definition of such objectives requires firstly specification of the components of driving competence which can and must be assessed within the framework of different forms of testing, and then, as a second step, detailed description of the competence levels and demand situations to which test questions are to refer. This task is yet to be accomplished, as is likewise empirical verification of the conceived models. To be able to perform such verification, it is necessary – as in the case of the theoretical driving test, for example – to use test results to model or restructure the test contents (see Chapter 3); with regard to the practical driving test, corresponding prerequisites are currently still to be established through the intended introduction of an electronic test report for the documentation of test results.

2.2 Draft of a model of driving competence

Findings on driving behaviour were already gathered in driving behaviour models in the 1980s. In the engineering and traffic sciences, such models often assist explanation of the interactions between driver and vehicle in motorised road traffic. Two of the most important fundamental approaches used to describe and analyse driving behaviour are Rasmussen’s three-level model of goal-oriented performance (1983) and the three-level model of driving demands by Donges (1982), which corresponds in turn with the three-level models of vehicle control by Bernotat (1970) and Michon (1985). In the widest sense, Rasmussen's model comprises notions of the attainment levels and acquisition stages of action competence, whereas Donges' model considers instead content-based demand aspects and the components of driving competence. Donges’ (2009) combination of the two models (see Fig. 3) appears particularly interesting for the present purpose, because the further development of different forms of testing and their positioning in a system of novice driver preparation require that both the content structure dimensions and the levels and stages of driving competence acquisition be taken into account.

Rasmussen (see left-hand side in Fig. 3) distinguishes different levels of behaviour control, and thus knowledge-based, rule-based and skill-based behaviour. These control levels are typified by different extents of automation in behaviour control. If the model is applied to driving behaviour, then it is characteristic for knowledge-based behaviour that the driver must consciously recognise, evaluate and interpret the demands of a traffic situation, in order to plan the appropriate next actions and the manner in which these actions are to be performed. In the case of rule-based behaviour, the driver applies one of a stored set of behaviour responses acquired through experience or practice, i.e. he already knows what is to be done as soon as a certain situation is recognised (in the sense of “if-then rules”); behaviour is effectively “semi-automated”. Skill-based behaviour, finally, is characterised by reflexive stimulus-reaction mechanisms (“routines”) which no longer require conscious control and are thus applied fully automatically. Automation of the necessary action sequences in certain prototypical traffic situations (e.g. changing road lanes or turning at a junction), realised on the basis of if-then rules and routines previously acquired and consolidated through ever greater practical driving experience, serves to relieve the driver's working memory. This releases working capacities for the handling of demands at a superordinate level (e.g. nav-

[10] “The stages of a competence model could also be interpreted as steps towards the acquisition of competence. In this case, a competence model would deliver statements on how the interactions of different elements of competence develop over the learning biography of children and adolescents, in other words how competence acquisition proceeds. These questions, however, are scientifically much more difficult to answer than that concerning levels of attainment within a particular age cohort of schoolchildren.” (KLIEME et al., 2007, p. 77)

igation), for auxiliary actions such as radio tuning, or for the mental processing of unexpected or new situations, including control of an appropriate reaction.

Viewed overall, the left-hand side of the model describes levels or acquisition processes in the automation of driving behaviour which corresponds to levels of driving competence (knowledge-based, rule-based, skill-based). Despite the fact that, strictly speaking, Rasmussen (1983) was here distinguishing types of behaviour, his model contains notions of a normative competence level model.

The right-hand side of the model, on the other hand, considers different content-based demands entailed by driving. These ideas could be useful for the elaboration of a structure model of driving competence. Donges (see right-hand side in Fig. 3) distinguishes three demand types contributing to driving behaviour: “Navigation”, “Manoeuvring” and “Stabilisation”. On the navigation level, the driver determines a destination and plans a driving route. The decision for a particular route must here take into account factors such as the expected duration, which may vary at different times of the day, the purpose of the journey, possible intermediate destinations and the safety of the route (e.g. the probability that the road has been gritted in winter). During the journey, it may furthermore become necessary to seek an alternative (e.g. in case of hold-ups on the planned route), which then involves re-orientation. On the manoeuvring level, the driver realises the planned journey; the chosen manner of driving is matched to the course of the road and the surrounding traffic. This requires safe performance of driving manoeuvres such as cornering, overtaking, lane changing or the negotiation of junctions. The task for the driver on the manoeuvring level is thus to observe the traffic, to maintain an appropriate driving line and to adapt the speed and distance to other vehicles to the situative circumstances. On the stabilisation level, finally, the driver makes those corrections which are required to avoid losing control of the vehicle. The reactions must be chosen correctly and applied to an appropriate degree.

It has already been mentioned that it would be desirable to combine content-based and graduated competence aspects in a single driving competence model. One interesting example in this direction is the overall view (see Fig. 3) in which Donges (2009) assigns his three content-based driving demands to the three behaviour control and automation levels defined by Rasmussen (1983). In the figurative sense, this yields a normative driving competence model indicating the automation (or competence) levels on which a driver should normally handle the various content-based demands of driving. It can be seen, for example, that it is hardly possible to automate navigation, and that this occurs mainly on the level of knowledge-based behaviour. This results not least from the fact that journey destinations, routes and road conditions will only seldom reoccur in identical combination; navigation is thus generally subject to conscious control and ties up mental capacities. By contrast, the handling (or manoeuvring) of a vehicle, which includes above all vehicle control, traffic observation, communication with other traffic participants, speed regulation and positioning of the vehicle in a given traffic situation, is accomplished more or less automatically by an experienced driver in the form of skill-based behaviour; it nevertheless requires also knowledge- and rule-based behaviour at times – above all in unusual or unexpected traffic situations. This means that vehicle handling can be automated to a certain degree, but must also be controlled consciously in part; a novice driver must first acquire the possible automation level through driving practice.

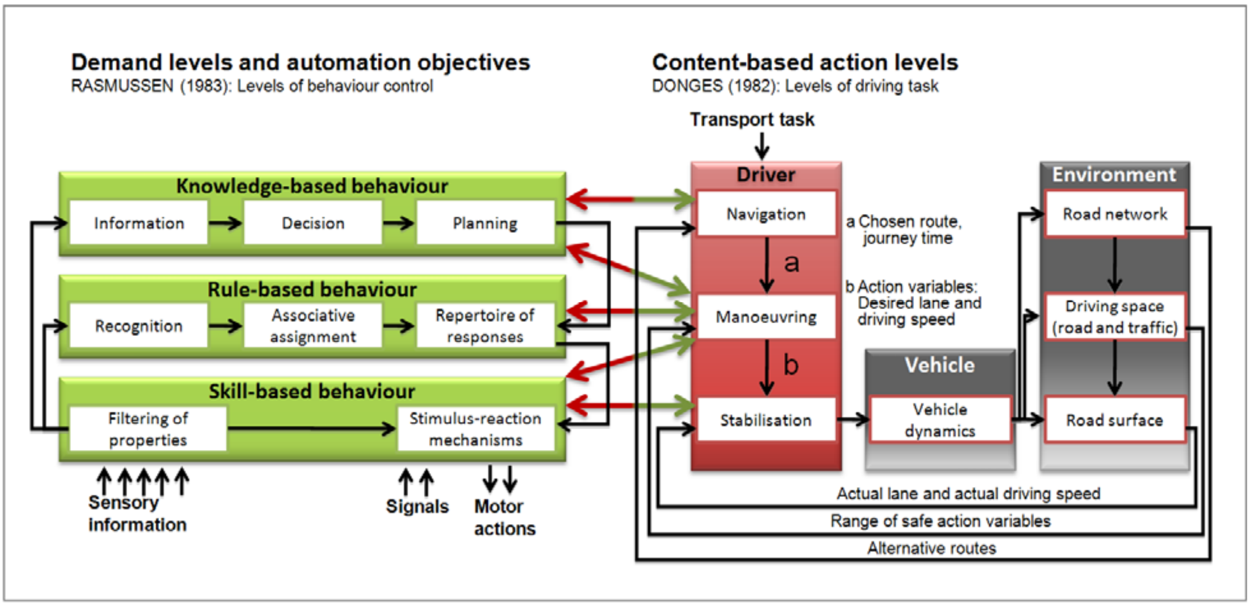

Fig. 3: Driving behaviour model according to Rasmussen (1983) and Donges (1982) (based on DONGES, 2009)

Behaviour on the stabilisation level, finally, is based on skills; it becomes automated over time and then only occupies a small proportion of working memory capacity.

Although not a driving competence model, the model presented by Donges (2009) offers a number of starting points for the elaboration of a driving competence model serving to systematise driving test contents. At the same time, two desiderata are brought to the fore: (1) Firstly, it appears expedient to supplement the model components representing the content-based demands of driving ("Navigation", "Manoeuvring" and "Stabilisation") with value- and norm-based components, in order to be able to map driving behaviour demands resulting from the social references of driving. (2) Secondly, the automation levels which originally served to explain the acquisition of skills convey a view of the psychic components, or the acquisition of driving competence, which is limited primarily to psychomotor aspects. It here appears more promising to relate the specific driving competence concepts to competence components and to avoid process references. The latter can also be deemed expedient because the action processes of driving are not all (necessarily) automated and doubts exist regarding the hierarchical process of driving competence acquisition (GRATTENTHALER, KRÜGER & SCHOCH, 2009). The manner in which these two desiderata were incorporated into the present conceptualisation of driving competence is explained in more detail below.

(1) Driving a motor vehicle is essentially a transport task which can be solved in very different ways. It is possible, for example, for a driver to navigate, manoeuvre and stabilise his vehicle in accordance with the demands, but nevertheless to endanger other road users or damage the environment. Traffic behaviour is thus influenced not only by navigation, manoeuvring and stabilisation demands, but also by the form of demand solution which the driver considers desirable and justified or which society prescribes. Social values and norms, and subsequently their acquisition and verification, are thus subjects for novice driver preparation. Following the frequently cited definition of Kluckhohn (1951), values are understood to be identity-building and implicitly or explicitly perception- and action-steering notions of what is justifiably desirable. They are acquired in the course of socialisation and identity formation, serve the individual as orientation in diverse situations, and ensure the stability of social systems. The latter also applies to norms: Norms are based on values, defined as behaviour demands for recurring situations and backed up by sanctions for the case of non-compliance; they thus represent the foundation of normality (LAMNEK, 2001).

The necessity to take values and norms, i.e. the social context of driving, into account in the contents of novice driver preparation already led Keskinen to define an extended driving competence model in 1996. His "hierarchical model of driving competence acquisition" integrates the aforementioned action control model by Rasmussen (1983) and the hierarchical risk model by van der Molen and Bötticher (1988), and supplements these two models by adding a value-related component. This thinking was later incorporated into the so-called "GADGET matrix"[11] (HATAKKA, KESKINEN, GREGERSEN & GLAD, 1999), in which the following content-based demand and action levels are distinguished:

1. "Vehicle manoeuvring" (lowest level: Control of vehicle speed, direction and position),
2. "Mastering of traffic situations" (adaptation to the demands of the current traffic situation),
3. "Goals and context of driving" (objectives, environment, social context, company in the vehicle) and
4. "Goals for life and skills for living" (highest level: Importance of the vehicle and driving for personal development).

Both the third action level and the newly introduced fourth level "Goals for life and skills for living" refer to the way a driver's personal value orientations and social attitudes influence his participation in motorised road traffic. One example of the demands of this level is the handling of influencing factors arising from the life circumstances of young novice drivers (e.g. a stronger quest for social acceptance and recognition). It is probable that for them, compared to older drivers, driving is more frequently associated with additional motives such as a striving for independence, the desire to impress and increased fun, which may lead to a reduced acceptance of traffic rules and to a greater readiness to take risks.

[11] "GADGET" stands for "Guarding Automobile Drivers through Guidance Education and Technology". The four described levels are combined with three vertical dimensions describing the essential training contents achieving safe participation in traffic. These dimensions are the driver's "Knowledge and skills", "Risk-increasing factors" and "Self-evaluation". The individual cells of the resultant 4 x 3 matrix represent a structural definition framework for driving competence and are used in countries such as Norway, Sweden and the Netherlands to determine the training contents and teaching/learning forms (and thus implicitly also the test contents and forms of testing) for novice driver preparation.

Assessing the four demand levels of driving behaviour according to Hatakka, Keskinen, Gregersen and Glad (1999) in respect of their suitability to describe the possible and necessary contents of different forms of testing, and comparing the GADGET matrix with the original three-dimensional demand models, the inclusion of a value level can be said to be an important enrichment. The combination of traditional navigation demands with the social context of driving and other value-related aspects on the third demand level, however, is suboptimal, as navigation is to be viewed as an significant independent demand and competence aspect: Inadequate navigation competence can quickly place excessive burdens on the novice driver on the manoeuvring and stabilisation levels, and it is for that reason that the systematic variation of navigation demands is among the contents examined in the driving licence test in many countries. Furthermore, it promotes the discrimination of a content-based demand model if the value- and norm-related demands are operationalised on a single (highest) demand level. The very subtle distinction between the largely situative "Goals of driving" (third level) and the essentially overarching "Goals for life and skills for living" (fourth level) may reflect pedagogical ambition with regard to the training objectives; it must be remembered, however, that conditions such as the licence candidate's driving or life goals play practically no role in the test situation. The corresponding competence prerequisites are thus only marginally verifiable, and that primarily in the form of factual knowledge.

The expansion of the traditional stabilisation level by Hatakka et al. (1999) to include the demands of direction control, speed regulation and vehicle positioning on a new "Vehicle manoeuvring" level also moves the focus away from the typical challenges facing a novice driver in the first phase of driving competence acquisition, and from a clear differentiation of test contents, particularly in respect of the practical driving test. The traditional stabilisation level addresses vehicle control processes which – for an experienced driver – are essentially automated and independent of specific driving manoeuvres; they must still be learned by the novice driver and are thus also to be taken into account separately during the driving test and in other learning progress assessments (for example under the aggravated driving conditions of special safety training). It is true that vehicle control is also maintained through regulation of the vehicle speed and position, but this must still be distinguished from targeted speed regulation and vehicle positioning in the service of driving manoeuvres (second demand level). Successful vehicle stabilisation imparts a sense of security to the driver (and naturally also to the driving instructor and driving test examiner), and thus establishes the operational and emotional prerequisites for completion of the driving manoeuvres necessary to master special traffic situations.

From our point of view, the arguments presented here demonstrate that, for driving licence testing, the relatively clearly discriminated demand levels "Stabilisation level", "Manoeuvring level", "Navigation level" and "Value level" provide a useful starting point for the determination of content-based driving demands, and thus also of content-based components of driving competence. A corresponding structure model thus starts with a basic operational demand level (the stabilisation level); this base supports a tactical level (the manoeuvring level, where the operational elements are arranged meaningfully into driving manoeuvres according to situative demands), followed by a strategic level (the navigation level where driving is planned) and an overarching value level. Consequently, these four demand levels should be used to systematise content for the different forms of testing and the different test items.

(2) Grattenthaler, Krüger and Schoch (2009, p. 82) treat driving competence as action knowledge which can be subdivided into the following three knowledge forms or acquisition levels:

1. "Explicit knowledge" in the form of situation prototypes and action scripts as central elements of top-down action planning (this knowledge is explicit in nature, can generally be reported and is thus accessible for verbal instruction),
2. "Implicit knowledge" in the form of motor schemata, further differentiated by feedback loops of action effect, environment perception (above all visually) and proprioception (this knowledge is generally implicit and unsuitable for reporting), and
3. "Process knowledge" as the integration of explicit and implicit knowledge (the integration is not sequential, but instead parallel, and requires prior successful resource control and self-evaluation).

This conceptualisation of driving competence as action knowledge corresponds with the initially described starting positions for the further development of novice driver preparation in educational psychology and with its theoretical roots in education and expertise research (BAUMERT, 1993; WEINERT, 1998). Driving competence as action knowledge thus embraces the facets knowledge, understanding, skill, ability, action and experience from Weinert's concept of competence. The facet of motivation, however, is in our opinion not taken into account adequately in this conceptualisation of driving competence and should thus be included explicitly alongside the three forms of knowledge in order to complete the picture: Process knowledge does enable the novice driver to master all demands associated with safe, environment-aware and energy-saving participation in motorised road traffic, but the driver must also be motivated accordingly.

One restriction to be added with regard to the driving test, however, is that the motivation to master driving demands appropriately can hardly be determined in the test situation, because all licence candidates aim to pass the test and will thus behave in conformity with the demands – possibly in contrast to their behaviour in later solo driving. Consequently, verification of the motivational prerequisites for driving competence is above all a task for medical-psychological assessments of the fitness to drive and not for the driving test.

If the thoughts presented under (1) and (2) are incorporated into a two-dimensional competence model combining content-based demands and the psychic prerequisites of driving behaviour, the result is Figure 4. It can here be assumed that both explicit (factual) knowledge and motivation are more relevant for the mastering of tasks on the higher demand level; implicit knowledge, on the other hand, is of greater significance for the mastering of demands on the lower levels.

The above model can now be used, for example, to classify more differentiated content-based components of driving competence and to structure the test items accordingly. At the same time, it permits the systematic elaboration of demand situations for the operationalisation of test content (see following section). Lastly, it is possible to describe the areas of content and driving competence covered by the different forms of testing. This, in turn, enables short-, medium- and long-term goals to be derived for the further development of different forms of testing in general and optimisation of the theoretical driving test in particular.

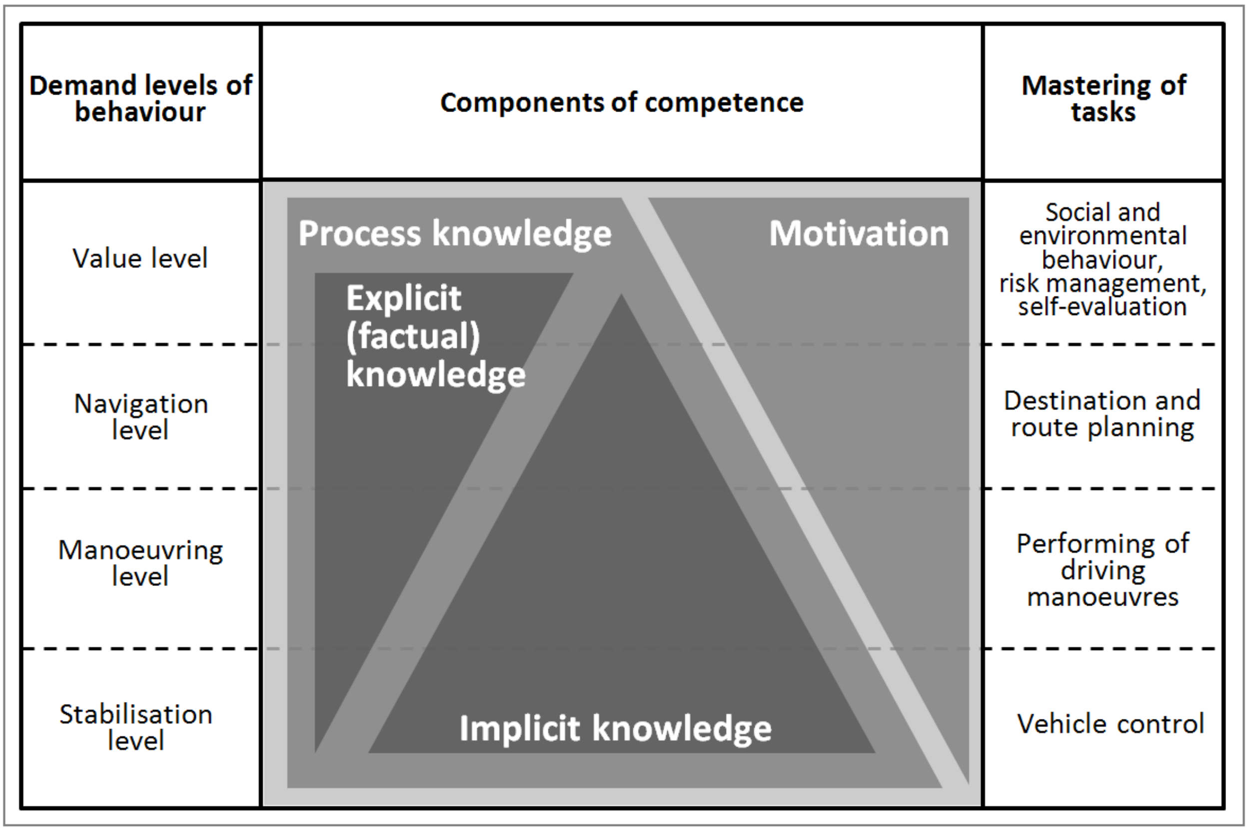

Fig. 4: Proposal for the structuring of content-based demands and the psychic components of driving competence

2.3 Description of demand situations

The driving competence model presented here is intended to aid decisions as to which content facets of driving competence are to be conveyed at which acquisition level during novice driver preparation and thus need to be assessed within the framework of the driving licence test. It is equally important, however, to identify the corresponding prototypical traffic situations which should be taken into account during novice driver preparation and must subsequently be reflected in the items contained in a driving licence test. This follows from the above theoretical considerations, according to which competences always relate to (demand) situations and should be operationalised for as many content facets of the training field as possible. Without definition of such demand situations, the driving competence model remains too abstract to support the concrete elaboration of test items. Consequently, this section describes the scientific basis upon which the TÜV | DEKRA arge tp 21 working group derived traffic situations for the operationalisation of test items and the principles applied when doing so.

Scientific basis for the operationalisation of driving competence in novice driver preparation
A starting point for the definition of demand situations to be addressed by test items is provided by the comprehensive analyses of demands applicable to motorised participation in road traffic by McKnight and Adams (1970 and 1972). The purpose of those analyses was to develop scientifically founded objectives and content for driver training. As the first step to this end, all possible or desirable actions relating to participation in motorised traffic were determined by survey and described in detail under the headings "Driver", "Vehicle", "Road", "Traffic" and "Environment". To facilitate structuring of the 1,700 driving behaviour components acquired in this way, they were then classified as either "off-road behaviour" or "on-road behaviour", depending on the general action context. Within these classifications, the individual elements of driving behaviour were grouped into so-called "driving tasks" according to either the action goal or the spatial or time-based components of the situation. The subdivision of "off-road behaviour" produced the following three categories:

- "Pre-driving tasks" (e.g. planning, loading),
- "Maintenance tasks" (e.g. routine checks, periodical technical inspection), and
- "Legal duties/responsibilities" (e.g. obligations following an accident, carrying of a driving licence and vehicle papers).

The elements of "on-road behaviour" were similarly assigned to one of three categories, each of which placed a focus on different situation references:

- "General driving" referred to behaviour which is required continuously during driving (e.g. observation, navigation).
- "Basic control" covered behaviour which serves primarily to control the vehicle (e.g. speed regulation, moving off).
- "Situational driving tasks" represented behaviour required as adaptation to specific traffic situations.

Taking up the aforementioned components of the road traffic system, the latter category of situational behaviour was further subdivided into (1) "Driving tasks relating to the traffic situation" (e.g. parking, overtaking), (2) "Driving tasks relating to road properties" (e.g. choice of lane, driving through bends), (3) "Driving tasks relating to the environment" (e.g. driving in particular weather conditions, night driving), and (4) "Driving tasks relating to the vehicle" (e.g. towing and being towed, behaviour in case of breakdown). A diversity of individual specific driving tasks was then assigned to each of these subcategories. To assess the road safety relevance of these driving tasks, finally, accident analyses and other empirical data were used to elaborate a system of (criticality) criteria, on the basis of which traffic experts were able to give their judgement of the safety relevance of individual driving tasks. The results of the demand analysis and criticality ranking were published in the form of a report (McKNIGHT & ADAMS, 1970). This report contained a detailed description of 45 driving tasks derived from the individual driving behaviour components, and specified a criticality index and corresponding criticality level for each element of driving behaviour.

Both the analyses of motorised traffic participation demands and the development, structuring and criticality assessment of driving tasks are recognised as exigent theoretical and methodical achievements on the part of McKnight and Adams. Consequently, their work has in the past been taken up repeatedly in Germany in connection with further development of the practical driving test (HAMPEL, 1977; STURZBECHER et al., 2010). And recently again, Sturzbecher, Mörl & Kaltenbaek (in press) referred to McKnight's and Adams' results in their definition and validation of driving tasks for a future optimised practical driving test, and in the elaboration of an electronic test report according to educational psychology principles.

With regard to further development of the theoretical driving test, it must be considered whether the categories of driving tasks and driving behaviour presented by McKnight and Adams (1970) could offer similar starting points for restructuring of the contents of the test items (see below). In accordance with the initially mentioned interaction of teaching, learning and forms of testing in the system of novice driver preparation, finally, it must also be asked whether the above studies could lend impetus to development of the system overall.

Analysis of accident data for Germany

One possibility to verify the road safety relevance of test items – as can already be derived from the example of driving tasks in McKnight and Adams (1970 and 1972) – is to establish a reference to German accident figures. Consequently, the TÜV | DEKRA arge tp 21 working group has conducted extensive accident analyses to identify demand situations which are of particular significance for the safety of novice drivers. These analyses are based above all on accident data acquired by the police in the German federal state of Saxony; the full results have been summarised in a corresponding interim report (WEIßE, 2011). A few examples from these results are presented here to illustrate the possible contribution of such accident analyses to the description of demand situations and at the same time to highlight the limitations.

Accidents recorded by the police are classified in various category systems, which are described in more detail in the aforementioned report. The category system used most frequently for analyses is that of “Accident type”, by which accidents are classified according to the situation in which they arise. As far as novice drivers are concerned, Figure 5 provides a general picture of their accident involvement.

It can be seen in Figure 5 that the accidents caused by novice drivers are most frequently so-called “Single-vehicle accidents”, which result from loss of control over the vehicle without the involvement of other traffic participants. One typical example is straying off the road in a bend, a situation in which the consequences of an accident are often serious due to the collision with trees. Second place is given to “Accidents in longitudinal traffic”, which occur as conflicts between traffic participants travelling in the same or opposite direction. Examples here are rear-end collisions or accidents when overtaking. The miscellaneous category “Other accidents” is followed by “Accidents at junctions/crossroads”, where the conflict arises between a turning or crossing vehicle with an obligation to wait and another vehicle with right of way at a junction or crossroads, or similarly at the exit from a car park or other property. The remaining categories are “Accidents when turning off”, which represent conflicts with other participants in traffic (including pedestrians) when turning off a road, “Accidents in stationary traffic”, where the accident is triggered by a conflict with parking vehicles, and “Road crossing accidents” caused by conflicts with pedestrians cross

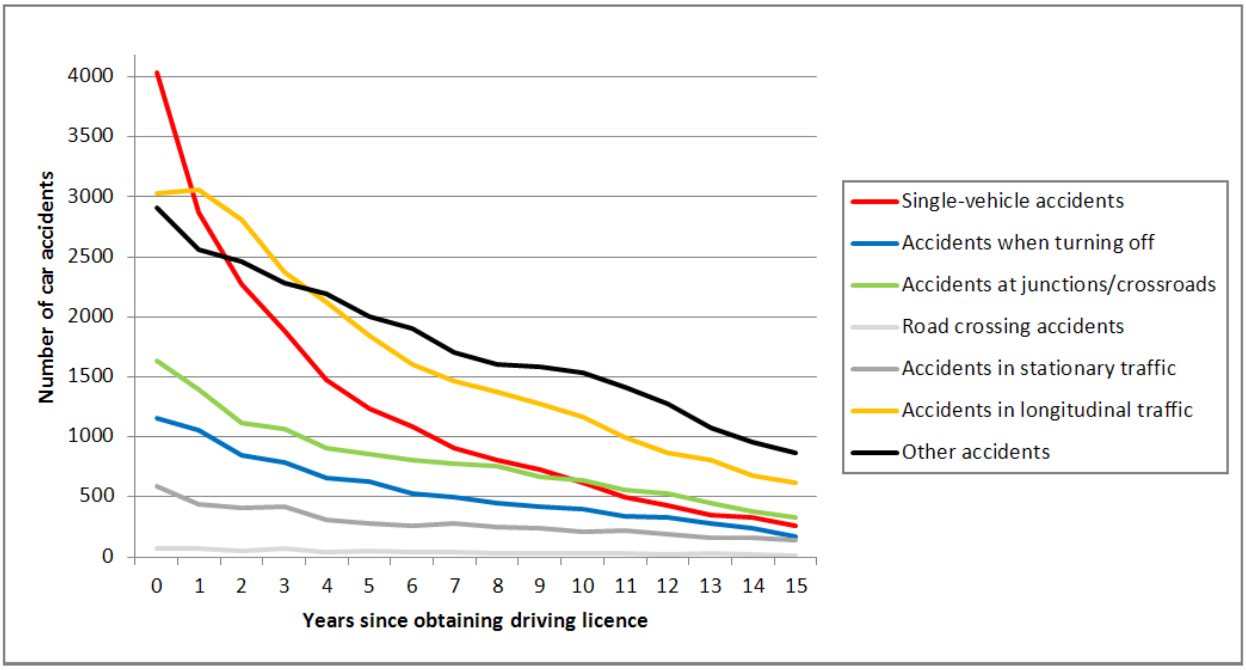

Fig. 5: Saxony 2004 to 2009 – Car accidents involving drivers who obtained a driving licence at age 18, grouped by accident type and duration of licence possession

ing the road. The particular ranking of novice driver accidents is attributable both to the frequency of corresponding trigger or demand situations in road traffic and to situation-specific competence deficits on the part of the novice drivers.

To be able to identify the novice-specific driving competence deficits, the ranking of accident causes for novice drivers was compared with the ranking for experienced drivers. The top category in the accident cause ranking for experienced drivers is “Other accidents”, followed by “Accidents in longitudinal traffic”, “Accidents at junctions/crossroads”, “Single-vehicle accidents”, “Accidents when turning off”, “Accidents in stationary traffic” and “Road crossing accidents”. This ranking differs from that of the novice drivers only with reference to “Single-vehicle accidents” and “Other accidents”, with the relative significance of “Single-vehicle accidents” declining as driving experience increases. This suggests that, above all in respect of “Single-vehicle accidents”, novice drivers are subject to situation-specific deficits which should be taken into special account in novice driver preparation and in the driving licence test.

This assumption is also confirmed by an alternative representation of the data showing the proportional decline in accidents in the individual categories with increasing driving experience (see Fig. 6).

The curve for the category “Single-vehicle accidents” also stands out from those of all other categories in Figure 6. The very similar curves of the remaining categories at the same time indicate that the different absolute figures reflect above all the frequency of exposure to the corresponding situations. Novice drivers are thus probably subject to a series of competence deficits which have similar effect in all these situations and are only reduced to the levels of experienced drivers over the course of several years. Evaluation of the other category systems in police accident records confirm this overall impression. Both the most frequent accident causes (inappropriate speed, inadequate distance to the preceding vehicle and failure to observe right of way) and the most frequent characteristics of the accident spot (bend, no special features, junction or crossroads) align well with the most frequent accident types.

A further outcome of these analyses was clear indication of the limitations of these category systems for the description of demand situations: For the most part, they cover only the stabilisation and manoeuvring levels of the described driving competence model. This is especially true for the example of “Single-vehicle accidents”. Viewed superficially, the accident analyses suggest that novice drivers have greater difficulty steering a vehicle safely around bends than coping with complex junction situations or overtaking manoeuvres. This conclusion, however, is inconsistent with the performances shown by novice drivers during driving school instruction and in the practical driving test. The reason for this discrepancy is probably to be found in the lack of information on important aspects of real demand situations from the value level. This includes, for example, the circumstances and purpose of the journey undertaken, which is not recorded in the police accident data, but may have led to the novice driver attempting to take a bend at a much higher speed than during driving lessons or the driving test. When deriving demand situations for novice driver preparation from accident analyses, therefore, it is imperative to consider carefully the model level to which the competence deficits leading to accidents belong.

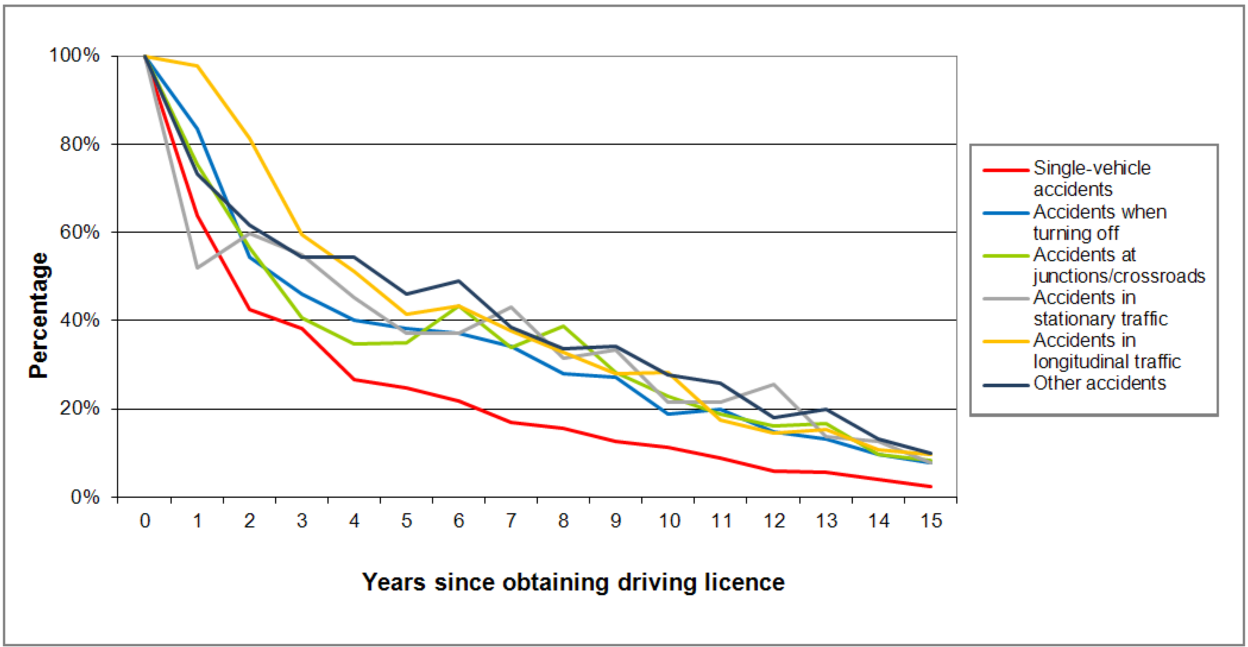

Fig. 6: Saxony 2005 – Accident frequency per year of driving career as proportions of the accident frequency in the first year of licence possession (driving licence obtained at age 18), grouped by accident type and duration of licence possession

It is also to be noted that the accident analyses of the TÜV | DEKRA arge tp 21 working group would gain in validity if further federal states were to follow the example of Saxony and make their accident data available.

2.4 Measurement of driving competence

The opening chapter presented the starting positions for elaboration of the innovation report for the period 2009 to 2010. It was also mentioned, that Genschow, Sturzbecher and Willmes-Lenz (in press) had analysed the systems of novice driver preparation in 44 countries, and – alongside learner assessments in the course of driver training – had on this basis distinguished three fundamental forms of testing, namely knowledge tests, traffic perception tests and practical driving tests. In this connection, the authors already pointed out that the contents of knowledge tests focus on factual knowledge, whereas action-related knowledge application often plays no particular role, and that the relatively new medium of computer-assisted traffic perception tests could probably improve the assessment of action knowledge. This initial statement can now be expanded and systematised with the aid of a number of fundamental considerations concerning driving competence acquisition and the presented competence model; further, future-oriented observations can then be found in the final chapter of this innovation report.

Fundamental considerations concerning driving competence acquisition in driving licence tests

If the driving licence test is understood as an instrument to measure driving competence, then its further development should aim to ensure objective, reliable, valid and not least also − from the point of view of test economy − efficient assessment of as many components of the previously developed driving competence structure model as possible in as many demand situations as possible. To illustrate these development requirements, the following “competence sets” can be distinguished on the abstract level:

- “Driving-relevant competences” (the set of all competences which can be assigned to the driving competence model and thus also represent relevant demands relating to participation in motorised road traffic) versus “Driving-irrelevant competences” (the set of all competences which cannot be assigned to the driving competence model and are thus not relevant for motorised traffic participation);
- “Conveyable driving competences” (the set of all driving-relevant competences which can be conveyed within the framework of novice driver preparation);
- “Assessable driving competences” (the set of all driving-relevant competences which, in principle, could be assessed within the framework of the driving licence test, irrespective of whether they are already assessed or not);
- “Assessed competences” (the set of all competences assessed within the framework of the driving licence test to date, comprising a subset of driving-relevant competences and a complementary subset of driving-irrelevant competences).

Figure 7 shows the direct competence sets and their relationships to each other. When interpreting these relationships between the competence sets, it must be taken into account that they are neither pre-established nor invariable. On the contrary: They shift in accordance with the results of research into driving competence acquisition; new findings with regard to the learning potential of individual training forms also play a role, as does the development of new possibilities to assess driving competence. It seems evident that the further development of novice driver preparation in general, and of the driving licence test in particular, is currently drawing greater inspiration and impetus from progress in the knowledge relating to driving competence acquisition and driving competence assessment, and not least from technical advances in the field of media for learning and testing, as opposed to changes in the overall road traffic situation, vehicle technologies or sociological characteristics of the target group of driving licence applicants. For future optimisation of the driving licence test, it is thus above all a question of determining which test contents are especially safety-relevant, how they can be assessed in a methodically optimum manner within the framework of special forms of testing, and at which points in the process of novice driver preparation these forms of testing must be employed, in coordination with corresponding teaching/learning forms and further forms of testing, to best fulfil their control and selection functions.

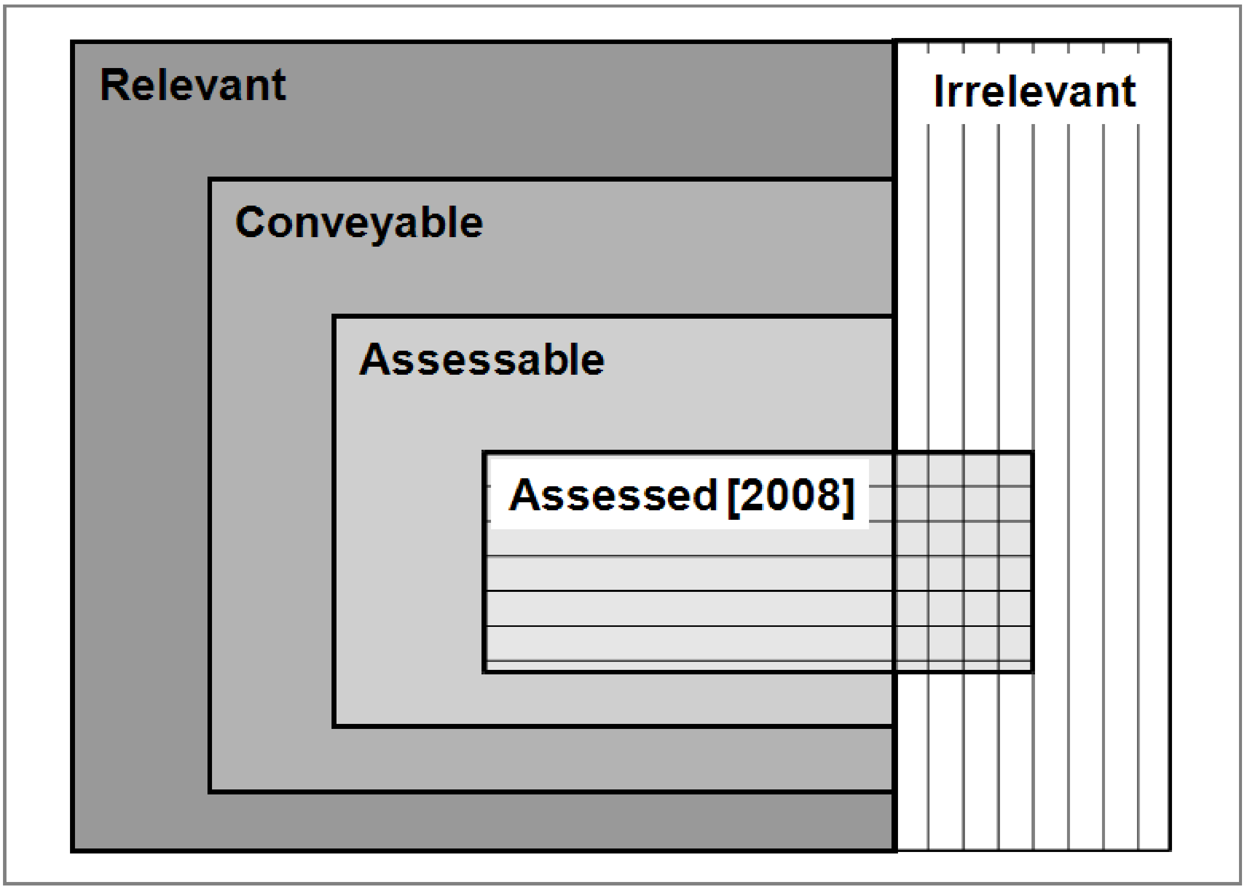

Fig. 7: Competence sets in the context of novice driver preparation and the driving licence test

The driving-relevant competences in Figure 7 result from the content-based demands of driving at the different action levels, as defined in the driving competence structure model (see Fig. 4), and from the described psychic components of competence. The question as to which of these competences can be conveyed within the framework of novice driver preparation is more difficult to answer than it would appear at first sight. Training practice to date indicates that explicit knowledge is conveyed predominantly in the theory classes and implicit knowledge in the course of practical driving instruction.

The integration of explicit and implicit knowledge into process knowledge is a lengthy process which begins in conjunction with driving school training and only reaches completion after several years of driving experience. British studies suggest that the attainment of a far-reaching automation of action sequences on the stabilisation and manoeuvring levels, which is typical for experienced drivers, could take up to three years, depending on the extent of driving practice (MAYCOCK & FORSYTH, 1997).

Alongside the traditional training elements of theory classes and practical driving instruction, new forms of learning, such as driving simulation training (including various forms of computer-assisted learning in which driving processes are simulated), accompanied driving, advanced courses and solo driving under protective regulations, are able to contribute significantly to the establishing of routine action sequences (GENSCHOW, STURZBECHER & WILLMES-LENZ, in press). To date, however, relatively few empirically based findings are available to show how the traditional and modern teaching and learning forms could be combined ideally to promote efficient driving competence acquisition (GRATTENTHALER, KRÜGER & SCHOCH, 2009).

In accordance with the Driving Licence Regulations (Fahrerlaubnisverordnung, FeV), driving licence applicants in Germany must demonstrate their possession of assessable driving competences by way of two traditional forms of testing, namely a knowledge test (the “theoretical driving test”, TDT) and a test drive (the “practical driving test”, PDT). The theoretical driving test can be described as a standardised and criterion-referenced knowledge test, which in its present form serves to demonstrate primarily declarative knowledge of rules and other facts (BÖNNINGER & STURZBECHER, 2005). The practical driving test, on the other hand, is from the psychological perspective a case of systematic behaviour observation and serves to demonstrate practical driving abilities and skills within the framework of a “work sample” (STURZBECHER, 2008 and 2010). Regarding the integration of the theoretical and practical tests into the overall system of novice driver preparation in Germany, it is stipulated that a candidate must have passed the theoretical test to be entitled to take the practical test; furthermore, the practical test must be taken

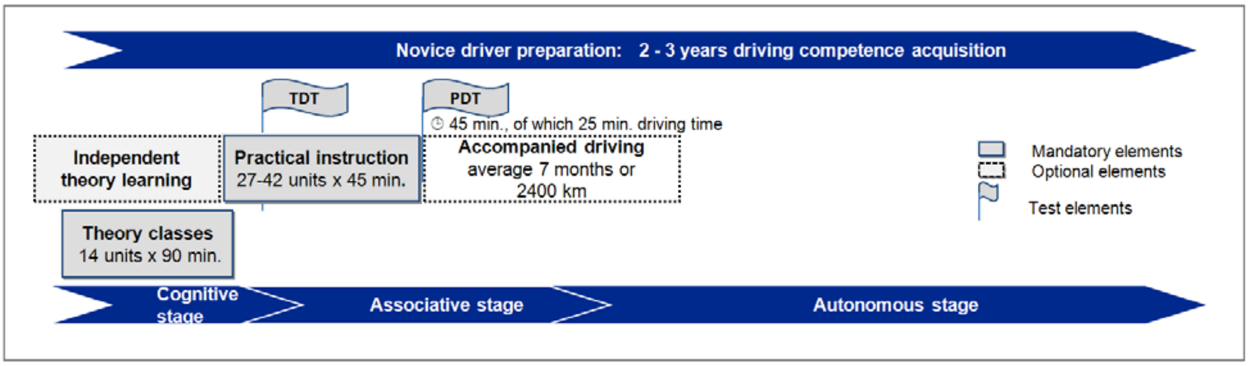

Fig. 8: Arrangement of the different forms of testing in the system of novice driver preparation in Germany (licence class B)

within one year after passing the theoretical test. The arrangement of the different forms of testing in the German system of novice driver preparation is shown in Figure 8.

In Figure 8, the aforementioned several-year period of driving competence acquisition is presented in slightly more detail in conformity with corresponding scientific studies (e.g. ANDERSON, 1982; LEUTNER, BRÜNKEN & WILLMES-LENZ, 2009; SHUELL, 1990). It starts with a “cognitive stage”, at which internal, primarily declarative knowledge of the actions required by participation in motorised traffic is accumulated by way of instruction and independent study. This is a prerequisite for the reception of further relevant information and for the ability to classify and process this information in individual knowledge structures. At the subsequent “associative stage”, the gathered stock of knowledge is then systematically improved, and thereby developed into implicit and finally action knowledge. The final “autonomous stage” serves perfection of the action knowledge; this enables ever more rapid and precise application of this knowledge with ever fewer mistakes, and not least to a reduction of the occupied attention and working resources.

The acquisition of action competence in general, and driving competence in particular, thus begins with the systematic development of flexible, connectable and transferable knowledge of the subject matter underlying the particular action, which in the present case means the circumstances of motorised road traffic (first step or phase). On this basis, it is then necessary to acquire the ability to apply the knowledge concerned effectively and in a manner appropriate to the situation, i.e. in the contexts of diverse traffic situations (second step). The final outcome of this process is the accumulation of a differentiated repertoire of problem - and situation-related action patterns, from which the immediately appropriate (driving) behaviour can be called up (third step). In the light of this three-step process of competence or expertise acquisition (ANDERSON, 2001; GREENO, COLLINS & RESNICK, 1996; GRUBER & MANDL, 1996), traditional knowledge tests, which address above all declarative knowledge and thus refer above all to the learning results of the first step, can be placed at a relatively early stage of the process of novice driver preparation, and indeed must be arranged so, if they are to provide meaningful motivation and support to knowledge acquisition. The greater operationalisation of procedural knowledge in further developed knowledge tests or new traffic perception tests, on the other hand, suggests placement of such forms of testing at a later point in the course of novice driver preparation due to the correspondingly longer learning processes (second and third steps). Practical driving tests serving to verify (elaborated) action knowledge, finally, must necessarily be placed at the end of novice driver preparation to be able to exploit their control effects.

If these correlations are taken into account, then it can be derived from Figure 8 that the early position of the theoretical driving test in the course of novice driver preparation, immediately after the basic theoretical instruction, i.e. before explicit knowledge has been transformed into implicit knowledge to any mentionable extent through driving experience or other (computer-based) realistic forms of learning, makes it difficult to assess traffic and hazard perception. The practical driving test is similarly taken at a relatively early point of the competence acquisition process, namely after approx. 1.5 to 3 months, i.e. at the beginning of the associative stage. Consequently, it permits assessment of only a minimum standard of driving competence, which is frequently still insufficient from the perspective of safe solo driving – as is indicated by the failure rate and the increased accident figures among novice drivers. The driving test thus controls only the initial phase of novice driver preparation; the level of competence which is possibly attained through accompanied driving (dependent also on the scope of such driving practice) remains unknown and thus irrelevant regarding the granting of a driving licence or provisions for solo driving under protective regulations.

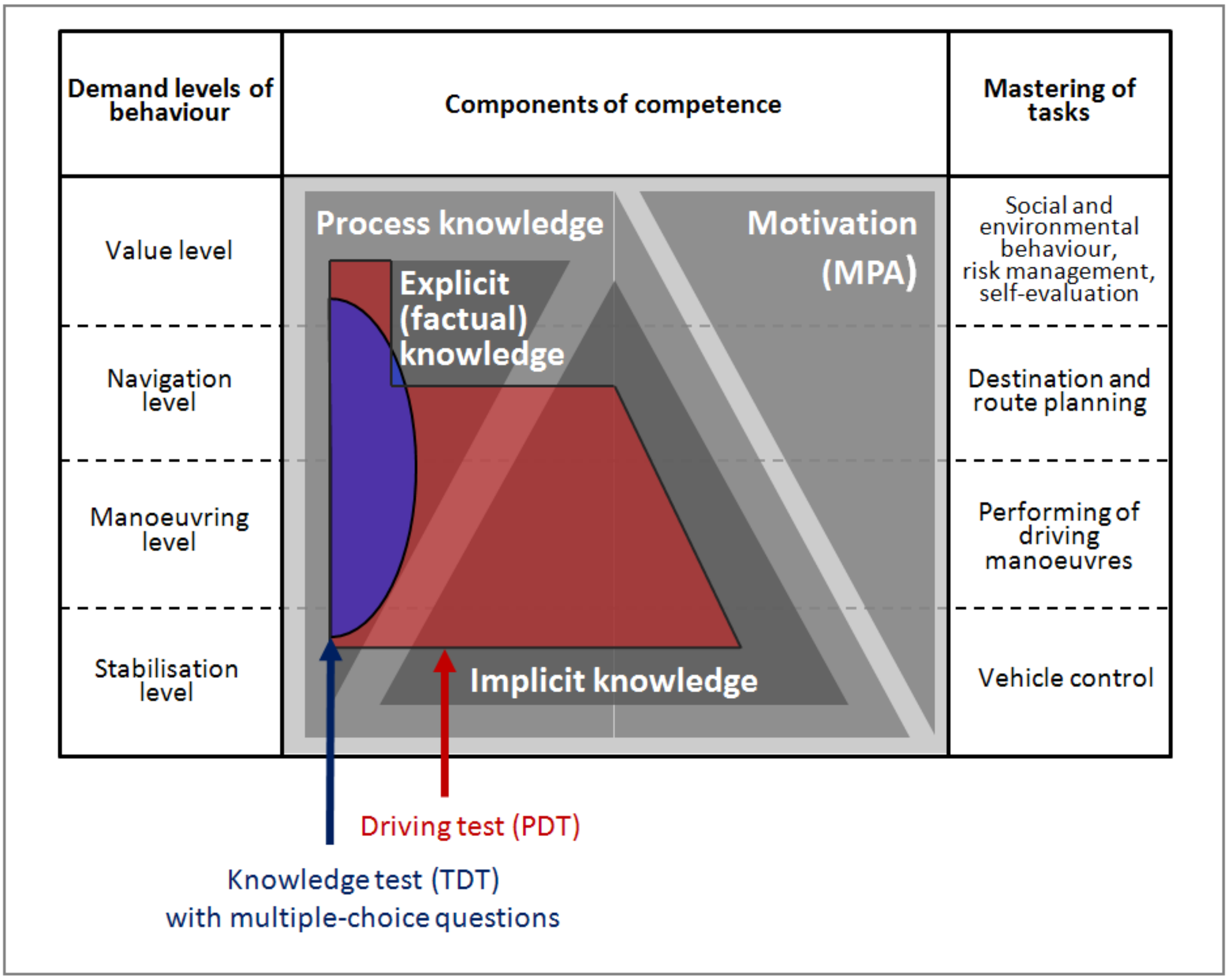

Fig. 9: Testing potential of the current forms of testing in Germany

The testing and development potential of the driving licence test

The described correlations between the forms of driving competence and knowledge on the one hand, and corresponding forms of testing on the other, can be illustrated by marking the areas of competence covered by the forms of testing in the competence structure model (see Fig. 4). The result can be seen in Figure 9, which indicates the overall testing potential of the current forms of testing in respect of the different driving competence components (see next page).

It can be recognised from Figure 9 that, above all with regard to the manoeuvring level (e.g. traffic observation, hazard recognition and hazard avoidance), there is no testing of implicit knowledge − i.e. of psychomotor schemata − outside the practical driving test, despite the fact that this would actually be possible at an early stage, for example within the framework of computer- or simulator-based tests. It would seem expedient to close this gap with innovative test formats such as traffic perception tests, in order to promote the acquisition of implicit knowledge – in the sense of the control function of licence testing – at early stages of novice driver preparation and, where possible, before participation in real traffic. This does not exclude also further optimisation of the assessment of factual knowledge within the framework of the theoretical driving test. Generally speaking, the task is to ensure that all assessable driving competences are indeed assessed validly with the various forms of future testing, and that the set of unnecessarily assessed competences is reduced to the greatest possible degree. In this context, the present innovation report focuses on the appropriate contribution of the theoretical driving test and on its optimisation and further development.

According to § 16 FeV, licence applicants must demonstrate in the theoretical driving test that they

- possesses adequate knowledge pertaining to the legal regulations decisive for the driving of a motor vehicle and to an environment-aware and energy-saving manner of driving, and
- are acquainted with the hazards of road traffic and with the behaviour necessary to avoid such hazards.

The first of these points concentrates evidently on the aforementioned explicit rule-related and factual knowledge on all demand levels of the described driving competence structure model − including the value level (e.g. environment awareness) and with a focus on the manoeuvring level. The second point already requires expansion of the targeted knowledge with regard to traffic situations and action sequences. This is thus no longer merely explicit knowledge, but rather knowledge of situation prototypes ("hazard situations in road traffic") and the corresponding action scripts ("behaviour to avoid hazards"), i.e. implicit knowledge. It must be noted, therefore, that the authors of the regulations expect not only the testing of factual knowledge, but also the testing of implicit knowledge within the framework of the theoretical driving test.

As already mentioned in Chapter 1, the TDT comprises parallel tests with 30 questions. Multiple-choice questions are the most frequently used format. In each case, an instruction or question is accompanied by two or three possible answers, of which at least one option is a correct answer. Further questions take the form of gap-fill questions, which require a numerical value to be entered at the placeholder. The text instructions to the test questions of the various parallel tests may also be illustrated by way of photographs or drawings in some cases. From the point of view of content, the test questions can be assigned to eight different subject areas: (1) Hazards, (2) Behaviour in road traffic, (3) Right of way, (4) Traffic signs, (5) Environment protection, (6) Regulations on the use of motor vehicles, (7) Technical aspects of a motor vehicle, and (8) Qualification and fitness to drive. It is here evident that the first two subject areas, in particular, also serve to assess safety-relevant implicit knowledge.

Over the past ten years, foundations have been laid for fundamental optimisation and further development of the TDT on a scientific basis (see Chapter 1). Many of the necessary optimisation measures have been implemented in the meantime; for example, computer-based tests have been used in all federal states since 2010. In the course of this restructuring, the existing pool of test questions for driving licence class B was subjected to scientific evaluation ("revision project"), with comprehensive statistical evaluation and content analysis. From the perspective of psychological testing, the analysis revealed the clear limitations of the current multiple-choice format, especially for the subject areas "hazard avoidance" and "affective-emotional behaviour in road traffic" (STURZBECHER, KASPER, BÖNNINGER & RÜDEL, 2008). These limitations arise essentially from fact that it is difficult to present the development and course of traffic situations, including the corresponding action sequences of those involved (driver, other traffic participants), in text descriptions and still images without overtaxing the reading competence and imaginative capacity of many candidates. Test questions with (extensive) textual instructions thus detract from the validity of the TDT, as they assess not only driving competence but also driving-irrelevant competences (see Fig. 7) such as reading competence and imaginative capacity; at the same time, test equality is reduced, as the nature of the test questions discriminates functional analphabets and those with reading weaknesses (on the basis of the PISA studies, a corresponding proportion of 15% may be assumed among the target group of the TDT; OECD, 2010). The current text-based instruction and response formats of the test questions, therefore, hardly permit the reasonable testing of implicit knowledge, particularly when it comes to specific knowledge in the subjects areas of hazard recognition and hazard avoidance, as prescribed by the Driving Licence Regulations (FeV). Consequently, experts have already been demanding fundamental changes in the methodical design of the questions in these subject areas for many years (HAMPEL, 1977); such changes, however, are only possible under the conditions offered by a computer-assisted TDT.

How could innovative test items and forms of testing for the assessment of implicit knowledge be designed? In the international field, it can be seen that a number of national systems of novice driver preparation – for example in Great Britain, the Netherlands and the Australian provinces New South Wales, Victoria and Queensland; further countries to follow – already use "hazard perception tests", either as an independent form of testing alongside a traditional knowledge test or as innovative, supplementary formats (GENSCHOW, STURZBECHER & WILLMES-LENZ, in press). This form of testing places the test categories traffic perception, hazard recognition and hazard avoidance in the foreground. Virtual traffic scenarios presented by the computer are the central medium serving to set the task, provide instructions and solicit a reply. The predominant test method is here to demand a correct reaction (e.g. the correct "driving decision" in the displayed scenario; at the same time, non-verbal response (e.g. the reaction time before a computer input) is also measured (see Chapter 5). The traditional knowledge test is thus experiencing further development and differentiation on the basis of the availability of new computer-assisted test media; with innovative test item formats and forms of testing, its weaknesses in the assessment of implicit knowledge, especially in the field of hazard avoidance, are increasingly being overcome. This also serves to bridge the aforementioned gap between the assessment of declarative knowledge in the TDT and the assessment of action knowledge in the practical test: The assessed driving competences are expanded with regard to implicit knowledge; with appropriate positioning of such forms of testing in the overall process of novice driver preparation, better utilisation of the control functions of driving licence

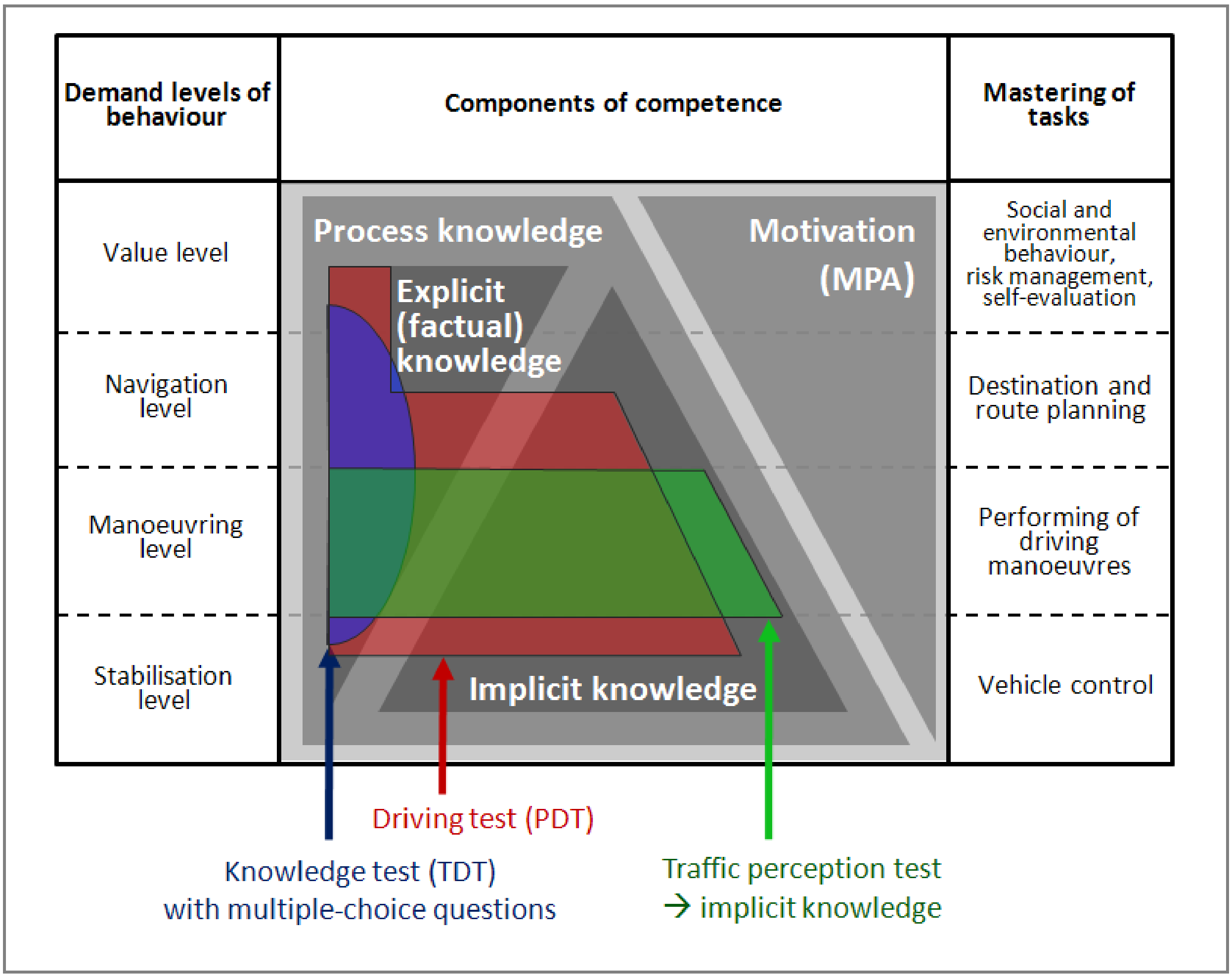

Fig. 10: Testing potential targeted for the individual forms of testing in Germany

testing is achieved. In Figure 10, the possible expansion of the competences assessed by the driving licence test after the incorporation of traffic perception tests is illustrated on the basis of the driving competence structure model (see Fig. 4).

From the theoretical perspective, and likewise in terms of content, the demands placed on innovative test item formats and forms of testing for the assessment of implicit knowledge in the field of driving competence can be further particularised with the aid of the information processing model (DODGE, 1982). This model was developed under the influence of the computer sciences and is used in cognitive developmental psychology, for example, to explain human behaviour in social situations; traffic situations can also be understood as social situations in which social actors interact against the background of behaviour-regulating values and norms. On this basis, innovative test items formats and forms of testing for the assessment of traffic perception, hazard recognition and hazard avoidance should address above all the following components of driving competence:

- Partial competences in the identification of relevant information and hazard cues in different traffic situations,
- Partial competences in the decoding and interpretation of different traffic situations (attention focussing and situation perception, targeted information searching on the basis of memorised situation schemata and taking into account the limited time available in a given traffic situation, information bundling, situation clarification),
- Partial competences in the clarification and weighing-up of action goals in different traffic situations (consideration of memorised reaction repertoires, retrieval of known reaction options from memory or construction of new reaction options),
- Partial competences in reaction planning and consequence anticipation, taking into account the specific contents and dynamics of different traffic situations (evaluation of reaction options on the basis of the expected consequences, conviction of the ability to perform the action and the appropriateness of the reaction, determination of the optimum reaction with a decision on further development of the original goal or selection of a new goal), and

- Partial competences in regulation of the appropriateness of reactions in different traffic situations.

Recourse to the information processing model (DODGE, 1982) effectively provides instructions for the operationalisation of test contents ("What can or should be assessed?") for innovative test item formats to assess implicit knowledge, above all on the manoeuvring level (see Fig. 4). One of the known benefits of the information processing model is that development is described as self-modification: Behaviour is modified on the basis of feedback, promoting the development of ever more effective information processing in the demand situations concerned (i.e. traffic situations in the present context). The model is especially suitable as a means to identify the individual processes which precede actual execution of a particular (traffic) behaviour response; it thus offers specific starting points both for the judgement of behaviour during the driving licence test and for teaching and learning methods aimed at driving competence acquisition.

To summarise: The test medium computer permits complete, differentiated and valid testing of the different components of driving competence to a greater degree than has been possible to date in driving licence tests. This applies equally for the traditional forms of testing – we can refer here to improved instructions for the multiple-choice questions of the theoretical driving test or to electronic test reports for the practical driving test – and for innovative forms of testing (e.g. traffic perception tests). The possibilities to assess driving competence components relating to traffic perception and to (rapid) hazard recognition and avoidance, in particular, are thus improved. Improved and coordinated assessment of the components of driving competence by different forms of testing, however, demands further elaboration of the construct of driving competence in respect of its content structures and acquisition processes, and strengthening of the research efforts addressing topics such as the trainability and measurability of driving competence components and novice-typical competence deficits. It is necessary, in particular, to pursue a wide-ranging review and discussion of the status of relevant scientific knowledge, and to test innovative approaches to the assessment of hazard recognition and avoidance competences and to the more comprehensive construct of "situation awareness"[12]. Existing scientific studies of the possibilities to acquire and assess abilities in hazard recognition, as well as findings relating to their concurrent and prognostic validity, for example with regard to the differing driving expertise of novices and experienced drivers or the later proving of novice drivers (overview: HORSWILL & McKENNA, 2004), indicate that more intensive occupation with this subject could be promising.

These theoretical considerations concerning driving competence acquisition establish a more precise basis for the tasks set in Chapter 1 for the work of the TÜV | DEKRA arge tp 21 working group during the report period 2009/2010, and at the same time offer contextual and organisational orientation for the future continuation of research and development work:

- The increased use of computer-generated images and dynamic driving scenarios in test item instructions is intended to optimise the traditional multiple-choice format, by reducing the dependence on the candidate's reading competence and by avoiding the more or less obvious solution hints in the present answer options. The emphasis of corresponding efforts to raise the challenge level posed by the testing with regard to content and methods are described together with the associated results in the following Chapters 3 (above all results of the summative evaluation of the TDT and their implementation) and 4 (above all the development of new instruction formats for the traditional multiple-choice questions, for example as a prerequisite for innovative test items and forms of testing).
- The objective of the further research and development work on innovative test items and forms of testing, which is to be intensified in the future, is to achieve a more realistic representation of actual road traffic circumstances in the TDT; this is achieved by testing implicit knowledge. It is very probable that this would also increase the validity of the TDT, as the candidate's actions during the test, subject to its particular conditions and options, would correspond much more closely to the driver's actions in reality (BÖNNINGER & STURZBECHER, 2005). Presentation of the initial investigations and research work in this direction forms the core of Chapter 5.

[12] Endsley (1995) describes situation awareness as "the perception of 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."

Jan Genschow, Andreas Krampe & Bernd Weiße

3 Evaluation and further development of test procedure

3.1 Background, objectives and potential of the TDT evaluation

Enhancement of the methodical quality of the theoretical driving test through continuous, scientifically founded evaluation was identified as a key objective for optimisation efforts in the final report of the research project “Optimisation of the driving test” (BÖNNINGER & STURZBECHER, 2005). Corresponding methods and procedures were developed within the framework of the subsequent revision and pilot project, and documented finally in the “System Manual on Driver Licensing (Theory Test)”. The goals and benefits associated with the introduction of a computer-assisted TDT, including its scientifically supported evaluation, were there presented in more detail, as follows:

1. Scientific basis for the TDT: Scientific foundations for the test content and methods enhance the quality of testing and improve the possibilities to distinguish reliably and validly between adequately and inadequately prepared novice drivers.
2. Evaluation of the TDT: The systematic and continuous evaluation of test results establishes prerequisites for monitoring of the effectiveness of driving licence testing in respect of the objectives.
3. Introduction of a quality assurance process: Process and result standards, including criteria for their fulfilment, are defined for the core areas of test content (e.g. methods for the elaboration, testing and approval of test items, and specifications for their corresponding assessment).
4. Prevention of schematic learning strategies: Measures such as the random ordering of test items and answer options prevent the application of schematic learning strategies and thus promote the candidate's motivation to analyse the underlying content intensively during preparation for the test.
5. Candidate-oriented aids: Measures such as spoken-language support for candidates with reading difficulties improve the test quality by excluding external influences on test performance.
6. Precautions against cheating: The introduction of computer-assisted tests implements effective measures against cheating (e.g. variation in the depiction of prescribed traffic situations, randomised order of possible answers, rotation of the questions in an initial sequence).
7. Improved feedback on test results to the licence applicant: Test reports with pointers to knowledge deficits improve the possibilities for a targeted appraisal of test performance by the candidate.
8. Continuous research work: The scientific basis of the TDT is strengthened through regular analysis and processing of the latest research in the relevant scientific disciplines. At the same time, own research projects are conducted (e.g. accident research studies, experimental studies with new formats).
9. Development of new test possibilities: Compared to a paper-and-pencil test, the use of a computer as test medium offers extended possibilities for the visualisation and presentation of test content and for the assessment of a broad spectrum of driving-relevant competences on the part of the candidate.
10. Continuous optimisation: Research work enables constant improvement of the methods and processes for realisation of an optimised TDT and corresponding quality assurance.

The introduction of a computer-assisted TDT in Germany implemented new methods and processes for continuous evaluation and further development in accordance with scientific standards. As a first optimisation step, the conclusions drawn from the evaluation results to date were used to revise the test items in the existing multiple-choice format.

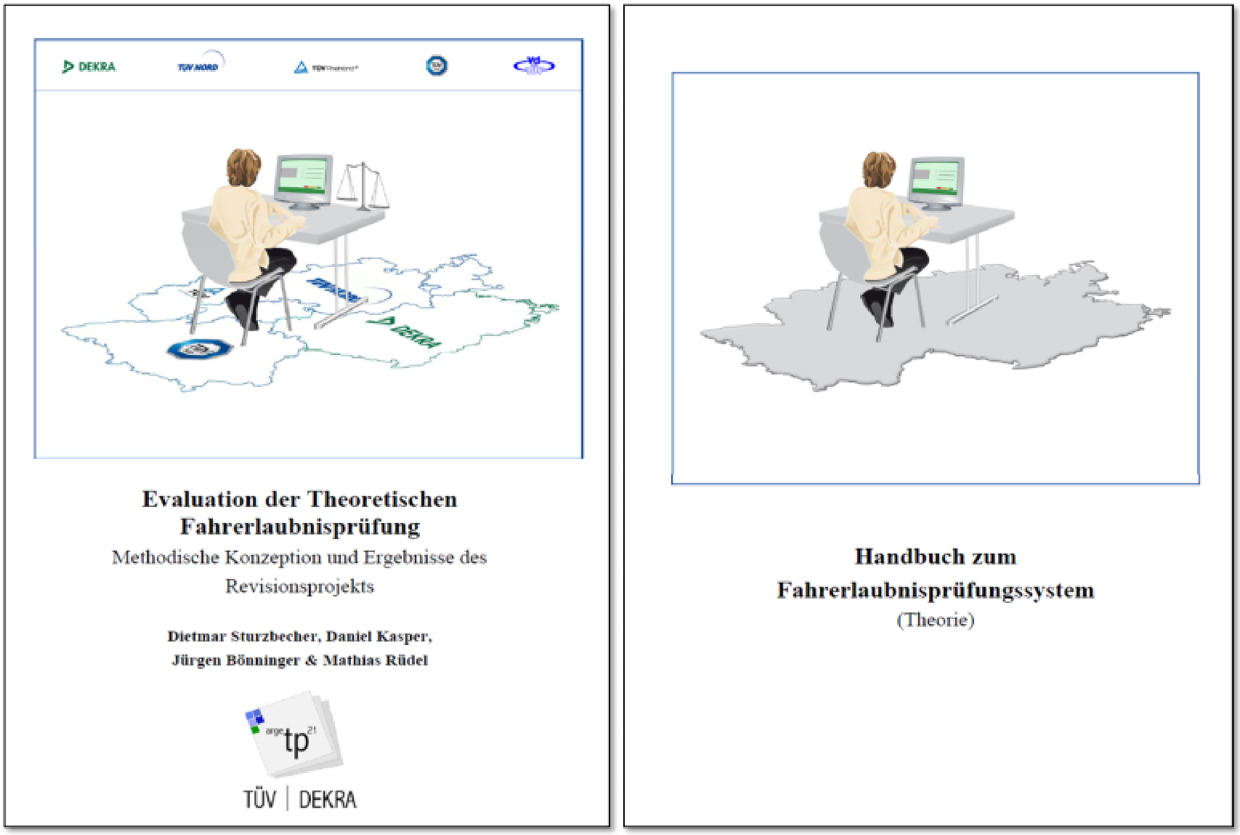

Fig. 11: Revision project report and system manual for the theoretical driving test

3.2 Development processes in the TDT evaluation

3.2.1 Further development of the methodical basis

The methodical concept for continuous evaluation of the TDT was developed and tested in its essential parts within the framework of the previously mentioned revision project (STURZBECHER, KASPER, BÖNNINGER & RÜDEL, 2008). Its validity has been proved and the concept anchored in the “System Manual on Driver Licensing (Theory Test)” (TÜV | DEKRA arge tp 21, 2008) as a fundamental component of the continuous evaluation and further development of the TDT.

The evaluation concept for the TDT provides for analysis and appraisal of the parallel tests and the test questions assigned to each test sheet. Analysis of the parallel tests serves to determine the test difficulty (as the mean proportion of correctly answered questions) and the variation of the proportion of correctly answered questions (indicating the extent to which candidates differ in respect of the proportion of correctly answered questions). To guarantee test equality in connection with the TDT, three empirically distinguishable forms of equivalence (SCHWENKMEZGER & HANK, 1993) are investigated: Psychometric equivalence, population-specific equivalence and experiential equivalence. For the analysis of test questions, the individual questions of a parallel test are first assigned to different category systems in accordance with the official catalogue of test questions, Annex II of the EU Directive on Driving Licences and the Learner Driver Training Ordinance (Fahrschüler-Ausbildungsordnung). In addition, they are described with regard to a series of formal criteria: For example, the official score penalties and the number of questions from the same category of the official catalogue are specified, along with explanation of the question type, the form of presentation and the form of treatment by the candidate. The methodical description of the questions includes specification of both the absolute guessing probability and the guessing probability after logical content evaluation[13]. The difficulty of a question (as the proportion of candidates who answer the question correctly) and its discrimination (as the correlation between the result for a single question

[13] The absolute guessing probability results from the number of all possible reply combinations; the guessing probability after logical content evaluation represents the number of all possible reply combinations minus the number of reply combinations which can be excluded after consideration of content aspects.

and the overall result obtained by the candidate for the whole test) are determined as statistical values.[14] Furthermore, the attractiveness and discrimination of the answer options are analysed. The psychometric analyses are supplemented with evaluative content analyses. The learning objective of each question is described and classified within traffic psychology models of driving behaviour and general systematics of driving competence. In addition, the challenge level of each question is considered, i.e. whether determination of the correct solution is dependent on specialist rather than general knowledge, for example. The relevance of the question contents for protective novice driver preparation is estimated by way of a comparison with the results of traffic science studies and accident analyses relating to novice-typical overtaxing and hazard situations. Finally, the reference between the question contents and the contents of driving school instruction are checked. The evaluation results are documented on the basis of prescribed report structures. A research report is produced for each parallel test, comprising the results of the parallel test analysis ("test sheet profile") and the question analyses of all questions belonging to the parallel test ("test question profiles").

The elaborated evaluation concept was tested empirically in the "revision project", by way of a systematic evaluation of approx. 20,000 theoretical driving tests completed throughout the country by test candidates wishing to obtain a first driving licence for vehicle class B (see Section 3.3). The revision project thus marks the beginning of continuous evaluation of the theoretical driving test, which has been conducted in its full extent since 2010. The evaluations of the test questions used in the TDT were continued with the evaluation of the test sheets for other licence classes.

3.2.2 Further development of the technical-organisational prerequisites

Questions which are identified as being in need of revision by the continuous evaluation must be replaced with optimised questions as quickly as possible. In the past, however, a very time-consuming voting process was necessary before such optimised questions could be published. The new methods and procedures described in the "System Manual on Driver Licensing (Theory Test)" are intended to accelerate this process. The new possibility for electronic voting here plays an important role: The time-consuming necessity for the involved parties to meet in person is avoided, which means that the approval process, with voting on the new test questions, can begin as soon as they have been elaborated.

As a technical prerequisite for this new voting arrangement, the TÜV | DEKRA arge tp 21 working group and the Federal Transport Ministry (BMVBS) have together developed a web-based platform, via which test question approval – i.e. acceptance of the question drafts by the federal states and subsequent approval by the transport ministry – can be accomplished in electronic form. This includes also the ongoing provision of documentation on the development of new questions or the modification of existing questions. The new process was presented to the Federal/Regional Expert Committee (BLFA-FE/FL) on 23rd/24th September 2009. The whole process chain for question approval is documented automatically and forms the basis for documentation of the overall process (see Fig. 12). A test run of this approval process was successfully undertaken and completed in the second half of 2009. For the federal state representatives responsible for the approval of test questions, contribution to the approval process is now much simpler to organise and occupies significantly less time than in the past.

A similar platform was made available to the "Theory test working group" responsible for test question revision at the VdTÜV, so that voting processes there can similarly be accelerated and documented comprehensively.

Furthermore, technical prerequisites have been established for publication of the current catalogue of questions at an Internet site. The corresponding application, however, is to date only available for the internal use of the Technical Examination Centres, as wider publication requires a prior policy decision of the BMVBS.

[14] The difficulty index and discrimination coefficients are evaluated by way of category systems. These category systems correspond to specific test-psychology-related characteristics of the theoretical driving test as a criterion-referenced test. They permit meaningful structuring of the questions in accordance with both the distributions of test results revealed during the revision project and the strategic goals of method development.

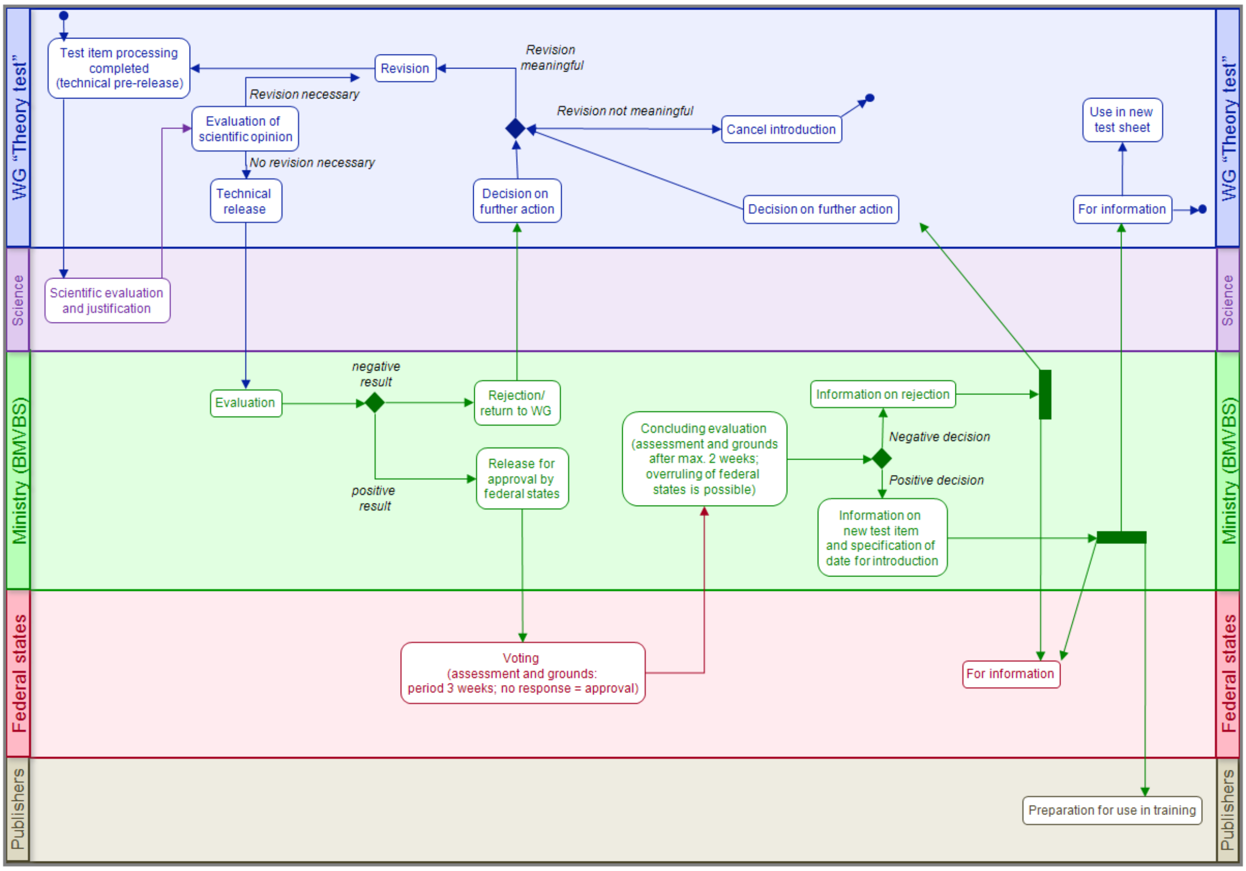

Fig. 12: Process of test question approval

3.3 Studies and measures within the framework of the TDT evaluation

The following section presents various studies which were conducted within the framework of continuous evaluation and further development of the TDT. These studies refer predominantly to research questions arising from the already realised transfer of conventional test sheets to the computer; another equally important topic of the presented studies is the concrete implementation of tasks anchored in the "System Manual on Driver Licensing (Theory Test)".

3.3.1 Evaluation of the test program interface

The introduction of a PC-assisted test system required evaluation of the program interface used to present and answer the test questions (see Fig. 13) from the perspective of the user, i.e. the driving test candidate. By way of an empirical study (HOFFMANN, BORNING & STURZBECHER, 2009), therefore, it was checked whether any usability deficits existed which could impair the candidate's test performance and thus diminish test equality.

The evaluation was realised in the form of a user survey. For construction of the questionnaire, suitable items were selected from the "IsoMetrics usability inventory" (GEDIGA & HAMBORG, 1999) – an established toolbox based on criteria for the usability of software interfaces in accordance with the standard DIN EN ISO 9241-110. The items were adapted linguistically to the specific characteristics of the test program interface and complemented with questions on overall satisfaction. Certain personal data describing the survey participants (age, gender, education, computer experience, test success) were also requested. The outcome was a questionnaire with which the standard criteria "Appropriateness of questions", "Self-explanation quality", "Controls", "Sensitivity to errors" and "Learnability" could be assessed and overall satisfaction with the program interface could be judged from the perspective of the user. The study sample comprised 199 driving licence applicants who took their theoretical driving test at the DEKRA branch office in Oranienburg (in the federal state of Brandenburg) between April and September 2009.

Fig. 13: The test program interface

The study results confirm the usability of the test program interface from the perspective of the candidate – 92 per cent of those surveyed expressed a high degree of satisfaction with the program. The assessment of usability was not dependent on the gender of the surveyed candidates, nor on their age, educational attainment or computer experience. The only factor which resulted in a reduced level of satisfaction with the program was failure to pass the test. It must be emphasised, however, that even among the subset of candidates who failed the test, more than two-thirds were satisfied with the test program. The lower satisfaction of certain unsuccessful candidates is presumably attributable to a self-serving bias effect placing blame for the failure on external influences (in this case the quality of the test program). Overall, the results of the study demonstrate that test equality is not impaired by limitations in the ergonomic design quality of the test program.

3.3.2 Continuous evaluation of test questions and parallel tests

On the basis of the evaluation performed within the framework of the revision project (STURZBECHER, KASPER, BÖNNINGER & RÜDEL, 2008), recommendations were elaborated with regard to content-specific and methodical optimisation of the parallel tests and individual test questions for licence class B. As a means to guarantee the psychometric equivalence of the parallel tests, a range was defined within which the marginally deviating levels of difficulty between the test sheets of the different parallel tests for a given licence class are to be considered equivalent ("difficulty corridor"). Where the difficulty of a parallel test was found to lie outside this difficulty corridor, it was recommended that the test sheet difficulty be adjusted through the targeted replacement of easy or difficult test questions, as appropriate. Such need for adjustment was determined for four parallel tests comprising a basic test sheet for initial granting of a licence, and for ten parallel tests representing the supplementary test sheet for licence class B. At the same time, it was suggested that the test questions from the subject areas "hazard avoidance" and "affective-emotional behaviour in road traffic", in particular, be improved both methodically and in respect of their content. These questions are described as significant for road safety by traffic experts, but usually displayed a high probability of correct solution, which was attributed to the often conspicuous nature of the question design (e.g. "solution hints").

Evaluation of the officially approved test questions and parallel tests has been continued on a regular basis since completion of the revision project by the Institute for Prevention and Road Safety (IPV

GmbH). Including the work done as part of the revision project, a total of 122 parallel tests in the test sheet categories “Basic sheet – Initial licence”, “Basic sheet – Extension”, “Supplementary sheet – Class B” and “Supplementary sheet – Class A” had been evaluated by the end of the report period for the present innovation report (2009/2010). A total of 768 different test questions were analysed. The corresponding sheet reports were communicated to the TÜV | DEKRA arge tp 21 working group.

Traffic-light function

At the 16th meeting of the development and evaluation group (EEG) on 17th June 2010, it was decided that, in addition to the previous summary assessment in text form, the determined necessity for revision of an evaluated test question was to be indicated by way of a three-tier colour-coded scale signifying the priority to be given to such revision (so-called “traffic-light function”). This change followed a suggestion from the experts of the VdTÜV working group responsible for further development of the theoretical driving test, who wished presentation of the results of the evaluation in a more condensed form, alongside the previous textual and tabular presentations, as an aid to prioritisation of the test questions in most urgent need of revision. The change was expected to achieve an improvement in the effectiveness and efficiency of evaluation processing.

The supplementary colour codes were applied for all test questions evaluated by IPV after completion of the revision project. For the report period of the present innovation report (2009/2010), this concerned a total of 463 different test questions from 9 further basic sheets for an initial licence (“versions”), 30 basic sheets for an extension, 11 further supplementary sheets for licence class B (“versions”) and 15 supplementary sheets for licence class A. A total of 61.1 per cent of these questions were marked “green”, which signifies no or no immediate need for revision. A further 35.2 per cent of the questions were classified “yellow”; these questions seemed to be in need of revision from the perspective of psychological testing, and revision should thus be considered and weighed up by the corresponding experts. The remaining 3.7 per cent of the questions were marked “red” – these questions displayed considerable deficits from the perspective of psychological testing and were thus to be excluded from further use in the TDT in their present form.

3.3.3 Studies concerning the randomisation of test questions and answer options

Since the introduction of the computer-assisted TDT at the beginning of 2008, the questions and answer options of parallel tests have been arranged in a random order, as a means to reduce the likelihood of manipulation of the TDT and to invalidate schematic learning strategies. In accordance with the underlying randomisation strategy, the order of the possible answers was varied randomly for all test questions. At the same time, in 2008 and 2009, the order of the questions covering the subject area “Behaviour in road traffic” at the beginning of the basic test sheet for an initial licence was also randomised.

The randomised ordering of questions and their possible answers may influence the test performances of candidates through so-called “order effects”. This would detract from test equality. Such order effects arise where the challenge level of a series of questions depends on the order of their presentation. To be able to exclude impairment of the test equality, possible correlations between the performance of the candidate and the order of the test questions and answer options were investigated empirically on the basis of the criteria “difficulty” and “discrimination” (STURZBECHER, KASPER & GENSCHOW, 2009). This analysis covered a total of 385 different test questions to be answered by the applicants for an initial class B licence and was based on test data records accumulated by the Technical Examination Centres over the period from April 2008 to April 2009. A total of 124,900 test data records were included in the analysis.

The analysis results showed that a statistically significant correlation between question difficulty and the order of the possible answers was present for only eight of the 385 test questions considered. In addition, two of the total of 30 basic test sheets for initial granting of a licence displayed a statistically significant correlation between the order of the test questions and the challenge level – though this affected only one question on each of the test sheets concerned.

As, overall, order effects were found for only a few of the test questions, and as these effects are for the most part attributable to methodical weaknesses or content correspondence, empirical proof can be assumed furnished for the methodical acceptability of the practised randomisation strategy. In case of broader use of the randomised ordering of test questions, however, similar verification should be performed for the subject areas concerned, in order to safeguard test equality. The results obtained to date indicate furthermore that corresponding studies could provide additional pointers to psychometric particularities in individual parallel tests and possibly to methodical weaknesses of the test questions.

The systematic acquisition of test data now enables empirical verification of the demanded methodical equivalence of the parallel tests (see Section 3.3.2). In addition, randomisation strategies providing for the randomised ordering of the test questions and answer options within a parallel test are being implemented (see above). The content structure of the test questions and the system for performance assessment have to date remained unchanged since introduction of the computer-assisted TDT. However, within the framework of optimisation of the TDT, starting points were identified for methodical and didactic improvements, and recommendations were elaborated for further development of the content structure; a corresponding overview is presented in the following sections.

3.3.4 Preparation for restructuring of the test contents

The content structure and the system for assessment of the "test sheets"[15] which are today used for the computer-assisted TDT are governed by the stipulations of the Driving Licence Regulations (FeV, Annex 7) and the Examination Guidelines (Prüfungsrichtlinie, PrüfRiLi, Annex 1). The test sheets are composed in accordance with the legal specifications on the basis of an official catalogue of test questions. They are divided into sections addressing basic and supplementary knowledge, and structured accordingly to cover the eight officially defined subject areas[16] (with further subcategories). In this system, each test question is always assigned to either the basic or supplementary section and to exactly one of the subject areas of the official catalogue. To produce the test sheet, a specified number of test questions is selected on the basis of test content (i.e. the assignment to a particular subject area within the official catalogue) and presentation form (illustrated or text question) in accordance with a distribution key laid down in the examination guidelines. For each driving licence class, the outcome is a set of parallel test sheets which are considered exchangeable. The equivalence of these parallel tests is verified empirically within the continuous evaluation of the TDT (see Section 3.3.2).

Deficiencies in the content structure of the test sheets are essentially attributable to existing weaknesses in the underlying system and structure of the test content (STURZBECHER & BÖNNINGER, 2005). These weaknesses result in part from the historical development of the driving test: The scope of the question catalogue and the content canon of the test questions have been widened successively to adapt the driving licence test to new legal regulations and changing traffic demands, albeit without systematic restructuring.

Subject area	Section in question catalogue	No. of questions	Total points
1. Basic content			
Hazards	1.1	8 (of which 4 illustrated)	32
Behaviour in road traffic	1.2	6 (of which 1 illustrated)	21
Right of way	1.3	3 (at least 2 illustrated)	15
Traffic signs	1.4	2 (at least 1 illustrated)	6
Environment protection	1.5	1	3
Total for basic subjects		20	77
2. Supplementary content	2.1 to 2.8	10	33
Total		**30**	**110**
Permissible score penalty 10; provided not two questions with weighting of 5 points answered incorrectly.			

Fig. 14: Selection and assessment of the test questions for driving licence classes A, A1, B, M, S, L and T (initial licence)

[15] The term "test sheet" is here used to mean the defined set of individual test questions to be answered by the candidate, as an analogy to the formerly printed paper sheet; these test questions are presented and answered on a computer, with the arrangement of the questions on the sheet being variable in order to prevent the use of forbidden aids (e.g. solution templates).

[16] The official subject matter is divided into eight subject areas: "Hazards", "Behaviour in road traffic", "Right of way", "Traffic signs", "Environment protection", "Regulations on the use of motor vehicle", "Technical aspects of a motor vehicle" and "Qualification and fitness to drive".

Under the classifications of the official question catalogue as currently valid, there is in parts a lack of clearly formulated distinction between the test subject areas (e.g. the subject area title “Behaviour in road traffic” relates to practically all aspects of road traffic). This means that the subject areas are not properly disjoint. The internal structures of the subject areas, furthermore, display very different degrees of differentiation. The aforementioned subject area “Behaviour in road traffic”, for example, comprises 30 subcategories. By contrast, there is no further content subdivision in subject areas such as “Right of way”, “Environment protection” and “Qualification and fitness to drive”. Finally, noticeable differences exist in respect of the classification system and terminology between the test contents as defined by the Examination Guidelines and the training contents specified by the Driving Licence Regulations. This structure of the test contents leads to test questions being equally capable of assignment to several subject areas from the expert perspective; unambiguous assignments are often realised solely for formal reasons. The verification of process and result standardisation for the procedure to elaborate test questions and test sheets should in future seek to ensure disjoint grouping of the test subject areas, which would avoid “overlapping” in the test contents.

The test contents should instead be restructured according to the principles of a professionally and scientifically founded, disjoint and exhaustive category system. This would overcome the existing distinction deficits in the internal structure of the test contents and the lack of discrimination in the assignment of test questions to subject areas; at the same time, prerequisites would be established for empirical confirmation of the validity of the test sheets in the course of the continuous evaluation of the TDT. If the restructuring of the test contents is to satisfy the above principles, then it is necessary to combine theory-based and empirical methods.

The theory-based approach entails analysis of the contents of relevant documents (Learner Driver Training Ordinance, test subject matter for the TDT, Annex II of the EU Directive on Driving Licences), in order to identify and describe the contents and structuring systems which constitute the relevant foundations of driving school training and driving licence testing. The analysed contents can then be classified logically to form superordinate content areas which are balanced with regard to the degree of differentiation and display no significant overlapping with the contents of other areas. In this context, the test contents can also be formulated as learning objectives, enabling test questions to be implemented as measurable quantities or test requirements (“The candidate knows, that …”). The theory-based approach to restructuring of the test contents should be flanked and underpinned by empirical studies on the basis of data obtained through the continuous evaluation of the test questions and parallel tests. The elaborated content structure for the test questions can then be analysed from different perspectives and modified, if necessary. The distribution of test questions between the subject areas, for example, is interesting for assessment of the extent to which the test subject areas are actually covered by test questions, and to reveal any imbalance in the question distribution between different subject areas. It is also conceivable to use factor analysis methods, to test the extent to which the theoretically developed content areas are reflected in the empirical data on candidate behaviour.[17]

3.3.5 Studies of the congruence of training and test contents

For a criterion-referenced test such as the TDT, full attainment of the learning objective by as many candidates as possible is considered the ideal case. The attainment of learning objectives can only be tested, if the test contents refer to training contents – the congruence of training and test contents is thus of great importance for the quality of a criterion-referenced test such as the TDT. Correspondingly, the “System Manual on Driver Licensing (Theory Test)” proposes expert assessment of the reference to training contents in each test question within the framework of the continuous evaluation, and documentation of the results in test question profiles (TÜV | DEKRA arge tp 21, 2008). Such systematic comparison of the training and contents permits identification of any missing references between training and driving test, and at the same time enables findings relating to the quality of individual test questions from the continuous evaluation to be taken into account. A guarantee for the congruence of training and test contents is thus not only imperative from the point of view of methodology, but also an important basis for targeted and systematic further development of the test questions.

[17] Factor analysis is used to check the factorial validity of a test. Proof is furnished by determining whether the questions of one subject area correlate more strongly with each other than those of different subject areas. Proof of factorial validity for the theoretical driving test is only possible if the test subject areas are structured disjointly. This path for empirical verification of the question content structure, however, is already hindered by the aforementioned methodical specifics of the TDT: As it represents a criterion-referenced test for which the candidate often prepares intensively, a relatively high proportion of the questions are answered correctly by most candidates. The result is a low variance of test performances, which limits the possibilities for structuring of the questions according to statistical properties (STURZBECHER, KASPER, BÖNNINGER & RÜDEL, 2008).

To provide a basis for decisions on future methodology in this area of the evaluation, topic area 5 “Right of way and traffic rules” of the Learner Driver Training Ordinance was selected for a representative comparative analysis of the training contents defined in the ordinance and the corresponding test contents of the driving test for a Class B licence. The analysis method and the analysis results were presented by the Institute for Prevention and Road Safety (IPV) at the 15th meeting of the development and evaluation group in Potsdam on 17th June 2009. The initiated systematic comparison is to be pursued further within the framework of the continuous evaluation. For the future, it will here be necessary to take into account both the necessary changes in the classification of test contents and, in the context of methodical implementation, the intention of a continuous systematic comparison with results processed and made available in the form of a database.

3.3.6 Variation of test sheet combinations

The revision project showed that the former practice of defining fixed pairs of basic and class-specific supplementary test sheets also offers potential for optimisation of the content structure of the test sheets (STURZBECHER, KASPER, BÖNNINGER & RÜDEL, 2008). This earlier pairing method was above all the result of practical requirements concerned with the printing of the sheets. At the beginning of the transition to a computer-assisted test, basic test sheets for an initial licence and supplementary test sheets for licence class B with the same sheet number were still combined into a single test. With this method, tests display extreme pass or failure rates whenever two test sheets which lie below or above the range of equivalence in terms of test difficulty (i.e. too easy or too difficult) are combined. It was thus recommended that the combinations of parallel basic test sheets for an initial licence and supplementary test sheets for licence class B be randomised in future. This led to individual very easy basic sheets being paired more often with a difficult supplementary sheet. Such random combination produces test sheets with increasingly similar average pass rates.

3.3.7 Changes to the system of assessment

In its current form, the TDT uses a system of weighted assessment: The test questions carry score penalties of two and five points, depending on their content and their significance for road safety, environment protection and energy saving. The weightings within the assessment system are determined on the basis of expert judgements. The score penalties assigned to each test question are published in the official catalogue of questions, and are also specified alongside the questions on the test sheet; they thus serve as a guideline for the candidate. When the candidate’s test performance is assessed, each question which is not answered correctly contributes to an overall result in accordance with its weighting. The test for initial granting of a licence for vehicle class B (basic test sheet and supplementary test sheet for licence class B) comprises 30 test questions with a total weighting of 110 points. The test is deemed to have been failed if a total penalty of ten points is exceeded, or if two questions weighted with a score penalty of five points each are not answered correctly. The weighting of the test questions must be taken into account when producing the test sheets, in order to guarantee an identical points total for all parallel tests.

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 (cf. LIENERT & RAATZ, 1998). Questions with greater significance with regard to traffic safety can still be awarded a higher number of points, but this possibility should be used sparingly, so as to avoid irregularities in the frequency distribution of the test results (BÖNNINGER & STURZBECHER, 2005). The revision project (STURZBECHER et al., 2008) sought empirical verification of the extent to which the present unequal weighting of the test questions is a suitable means to distinguish between successful and unsuccessful candidates. The results revealed that there is no statistically significant correlation between the current score penalty assignments and test success. The decisive factor for TDT success is thus not which questions a candidate fails to answer correctly, but how many. This finding is also valid for the so-called “ten-point rule”, which acts as a kind of “make-or-break criterion”. The ten-point rule stipulates that a candidate with a total score penalty of ten points still fails the test, if those points accrue from two incorrectly answered questions weighted with a score penalty of five points each; this applies to all questions from the subject area “Right of way” in the basic section to be answered for all vehicle classes. The analyses show that only very few candidates fail the test under the ten-point rule.

On the basis of the available empirical results from the revision project, it can be concluded that the present, differentiated assessment system makes no contribution to the selection function of the TDT and thus to the identification of those candidates who fail to meet the content-referenced minimum requirements of the test. This is not the case, however, with regard to the control function of the theoretical driving test: The findings in respect of the ten-point rule can be deemed evidence that candidates prepare particularly intensively for questions from a “make-or-break” subject area. The consistently high weighting of test questions in a particular subject area, in combination with a make-or-break criterion, is thus capable of exerting a positive influence on the learning behaviour of driving licence applicants. The sparing and pedagogically targeted use of weightings is a suitable means to encourage desirable learning effects on the part of the candidates.

The results of the revision project enable recommendations to be formulated for simplification of the present assessment system. One proposal is that the test questions should be assigned equal significance as far as possible. If the current emphasis on the subject area “Right of way” is to be maintained, these test questions could be given a higher weighting (two points, for example). Such an assessment approach would simplify the present system without introducing fundamental methodical changes. It should also be considered whether targeted “make-or-break criteria” could be used in place of different weightings to provide pedagogically meaningful learning motivation and to promote the candidates’ acquisition of particularly safety-relevant knowledge. A weighted make-or-break criterion could serve, for example, to ensure the communication and acquisition of new knowledge which is considered necessary in the context of traffic safety, and to offset typical competence deficits among driving licence applicants and novice drivers, should such deficits emerge from the evaluation of test questions or accident analyses.

3.4 Conclusions from the TDT evaluation

The studies to date show clearly that the now computer-assisted TDT displays an appropriate level of quality in terms of its practical usability and the safeguarding of test equality. Building upon the results of the revision project, the continuous evaluation of test questions and parallel tests at the same time demonstrates that, in principle, the majority of test questions in use fulfils the intended function of competence verification in accordance with various criteria relating to psychological testing. The methodical weaknesses which are by all means present in the test questions are eliminated or reduced through systematic methodical and content revision. Assessments from the perspective of psychological testing, documented in test question profiles on the basis of a corresponding evaluation, provide founded recommendations for further action. Nevertheless, the statistical and psychologically oriented analyses conducted so far are unmistakable pointers to the limitations of the present test question formats, especially for the subject areas “hazard avoidance” and “affective-emotional behaviour in road traffic”.

When considering question revision and optimisation, a distinction must be made between revisions which can be achieved by changes within the existing question format, and those which go beyond the scope of the multiple-choice and gap-fill questions used to date and are thus dependent on new, innovative formats. Newly developed question designs which combine the multiple-choice question format with dynamic driving scenarios to depict the relevant traffic situations are here an important step towards further improvement of the test quality, but they are still not sufficient to overcome the existing methodical limitations of the multiple-choice question format for the assessment of hazard recognition and hazard avoidance.

With the introduction of a computer-assisted TDT, the conventional test design was simply transferred to the computer, without changes to the content structure or assessment system of the test sheets and test questions. Within the framework of work on optimisation of the TDT, however, various starting points for methodical improvements were developed; they are summarised in the following overview. Three main points of interest were identified: (1) The methodical thoughts outlined here with regard to restructuring of the test contents must be developed further and tested empirically. The expected result is a theoretically founded and empirically confirmed classification of test contents, which would satisfy the principles of a scientifically founded, disjoint and exhaustive category system. Such systematic classification would establish the prerequisites for assessment of the validity of the test questions used in the TDT and for continuous empirical verification and optimisation in the course of future evaluation. (2) Restructuring of the test contents should be accompanied by optimisation of the currently practised system for the assessment of test performances. It is here proposed to significantly reduce or even abandon the presently prominent weighting of test questions by way of different score penalty assignments. It should be considered whether the sparing or temporary use of weightings, or even better the application of make-or-break criteria to certain test questions are suitable means to ensure the communication and acquisition of new knowledge which is considered particularly necessary in the context of traffic safety. (3) Last but not least, systematic analysis of the test contents is an important starting point for work to enhance the content validity of the TDT. This refers on the one hand to the congruence of training and test contents, as is generally to be demanded for criterion-referenced tests; given the availability of computer-assisted learning media and innovative presentation forms for the depiction of traffic-related situations in novice driver training, the possibilities for continuous updating of the comparisons between training and test contents should be considered. At the same time, increased content validity can also be expected if the work on the development of test questions continues to incorporate the findings of traffic science studies and accident analyses.

3.5 Implemented steps towards further development of the TDT

The positive results of a survey to gauge user satisfaction indicated that the test program interface developed to implement the computer-assisted TDT could be kept unchanged. To prevent test manipulation, the possible answers to the individual questions are presented in random order, the basic and supplementary test sheets are no longer combined as fixed pairs, and the order of initially the first six and later all test questions was similarly randomised.

The need for revision of the questions analysed to date was evaluated against educational and test psychology principles, from the linguistic perspective and on the basis of expert opinions. All three assessments contribute to a final decision on the need for revision by the VdTÜV working group responsible for further development of the theoretical driving test, and a corresponding recommendation to the BMVBS and the representatives of the federal states. To improve the efficiency of such preparatory evaluation by the working group, a “traffic-light function” was introduced during the report period to signify the priority for revision revealed by the evaluation results. To this end, the TÜV | DEKRA arge tp 21 working group developed an evaluation tool similar to the question approval tool for the BMVBS.

The recommendations pertaining to restructuring of the overall catalogue of questions and to the simplification of question weighting have not been implemented to date. They must first be formulated in more concrete terms, taking into account the legal aspects of driver licensing, as a basis for discussion by the branch experts and the issuing of recommendations for specific amendments to the corresponding regulations and guidelines.

Tino Friedel, Bernd Weiße, Jan Genschow & Andreas Schmidt

4 Development of innovative test formats with computer-generated scenarios

4.1 Background, objectives and potential for the use of computer-generated scenarios

The transition to the use of computers as test medium for the TDT opened the door to completely new possibilities for test question design. Applying the question categories defined for the field of psychological testing by Lienert and Raatz (1998), the test questions of the TDT can be classified by question type, e.g. multiple-choice, assignment, sorting and gap-fill questions (BÖNNINGER & STURZBECHER, 2005). Each instance comprises an element to present the actual question (“instructions”) and an element to acquire the candidate's answer. The various design possibilities for the instructions (including illustration of the underlying traffic situations) are here referred to as “instruction formats”, and the different designs for answer templates as “answer formats”. This deviates from the terminology used by Malone, Biermann, Buch and Brünken (2011), who used the designation “presentation format” instead of instruction format; this is considered ambiguous, however, as both the answer templates and instructions are presented in different designs and formats.

In the past, the TDT comprised mainly multiple-choice questions, for which the instructions and the answer template or answer options were presented in text form (the work to expand the spectrum of available forms of testing − e.g. knowledge test, traffic perception tests − is described in the subsequent Chapter 5, together with the associated item types and display formats. Presentation has to date been limited to textual descriptions, photographs and simple graphics. The switch from paper sheets to a computer test medium now permits the use of more diverse display formats, as it is no longer necessary to print the test sheets. For the most part, it is also possible to maintain the present multiple-choice question format, which provides for a relatively simple transition process for the organisations involved. The development of innovative visual illustration possibilities to expand the range of display formats, in particular instruction formats, thus represents the next step towards optimisation of the TDT. The work undertaken to this end is described in the present chapter.

Expansion of the instruction formats is to be accomplished in two steps. The first step is to replace the photographs and diagrams used to date with computer-generated images. This brings the following benefits:

- In the past, considerable effort and expenditure was required to produce or update the photographs used, as the relevant situations and traffic participant constellations had to be found, or else simulated, in the real traffic environment. To avoid this revision outlay, the photographs remained in the question catalogue over very long periods, which gradually impaired the attractiveness of the depiction and thus acceptance of the questions. Computer-generated illustrations, on the other hand, can be created and updated efficiently. This guarantees a constantly modern appearance and a high level of acceptance. Furthermore, the complexity of the depicted traffic situation can be tailored to the objective of the question concerned.
- The decisive advantage of computer-generated images over photographs is the facility for efficient variation of the image content. Changing of the displayed objects or traffic environment (without altering the fundamental content of the traffic situation) is able to prevent schematic memorisation. The candidate does not immediately recognise a particular image during the test and must thus consider the contents of the depicted traffic situation actively to determine the correct answer to the question.

Fig. 15: Example of computer-generated presentation with variants

As a second step, the present catalogue of TDT questions is to be supplemented by questions which use computer-generated video sequences to illustrate the question instructions or the underlying traffic situation. This offers the following additional benefits, in addition to those already mentioned above:

- It is frequently necessary to add textual descriptions to static illustrations in order to depict more complex traffic situations or to explain the development of a situation over time; this raises the demands placed on reading competence and may provide solution hints. Dynamic illustrations, on the other hand, permit self-explanatory presentations of even complex traffic situations and event sequences. This achieves a greater correspondence between the demands to be met to solve the test question and the appropriate real situations. It is thus possible, for example, to forego technical terminology previously used in training, which is often a source of solution hints.
- Furthermore, a dynamic format permits the incorporation of hazard cues during the course of situation development, without these cues necessarily being visible in the final image (e.g. a cyclist in a mirror blind spot or children hidden by parked vehicles). In this way, it is possible to assess not only the candidate's plain judgement of the hazard potential of cues via a static illustration (e.g. “Do children playing at the roadside constitute a hazard?”), but also whether the corresponding hazard cues are anticipated, recognised and taken into account in the chosen actions over the course of a situation. Aspects of driving competence such as situation awareness and hazard recognition can thus be tested in a diversity of forms, and the validity of the questions is also raised.

The development of traffic situations could naturally be presented by way of real videos. Compared to computer-generated scenarios, however, they share the same disadvantages as the photographs and graphics used to date (i.e. the high costs of creation, updating and the elaboration of variants).

4.2 Development processes for the elaboration of instruction formats with computer-generated illustration

4.2.1 Elaboration of software for the creation of computer-generated images – VICOM

Before use could be made of the aforementioned benefits of computer-generated illustration, it was necessary to develop a software tool with which the corresponding traffic scenarios can be created, adapted and varied. The level of the demands placed on users of the software (e.g. IT knowledge, prior experience with 3D programs) was to be kept as low as possible, in order to maintain accessibility for as large a circle of traffic experts as possible. On the basis of these specifications, the software VICOM (Visual Components) was developed by the TÜV | DEKRA arge tp 21 working group. The development of such complex software requires a multitude of decisions. A comprehensive presentation of all the associated decision processes (on aspects including object libraries, user interface, detail resolution, variant generation, colours, contrast, angle of view, etc.) would far exceed the scope of the present report. A number of fundamental decisions with direct impact on the form in which traffic scenarios are to be presented in the TDT are nevertheless to be mentioned briefly in the following. More detailed documentation of these and all other decisions can be found in the reference document (TÜV | DEKRA arge tp 21, 2011) on optimisation of the TDT.

To maximise the validity of test questions, the demands to be handled to interpret the depicted situations should coincide as far as possible with those applicable in the corresponding real traffic environment. In the latter, the driver is surrounded in all directions by constantly changing visual cues. A sharp image, however, is only seen at the centre of gaze. Nevertheless, objects in the driver's peripheral vision can at least be identified and selected as a target for the next saccade (JOOS, RÖTTING & VELICHKOVSKY, 2003). With a series of head and eye movements, the driver thus “scans” the visual environment ahead of him. The rearward visual environment is observed on a situation-specific basis, usually with the aid of rear-view mirrors, which are similarly located in a forward position. The horizontal extent of the continuously observed environment is illustrated in Figures 16 A and C.

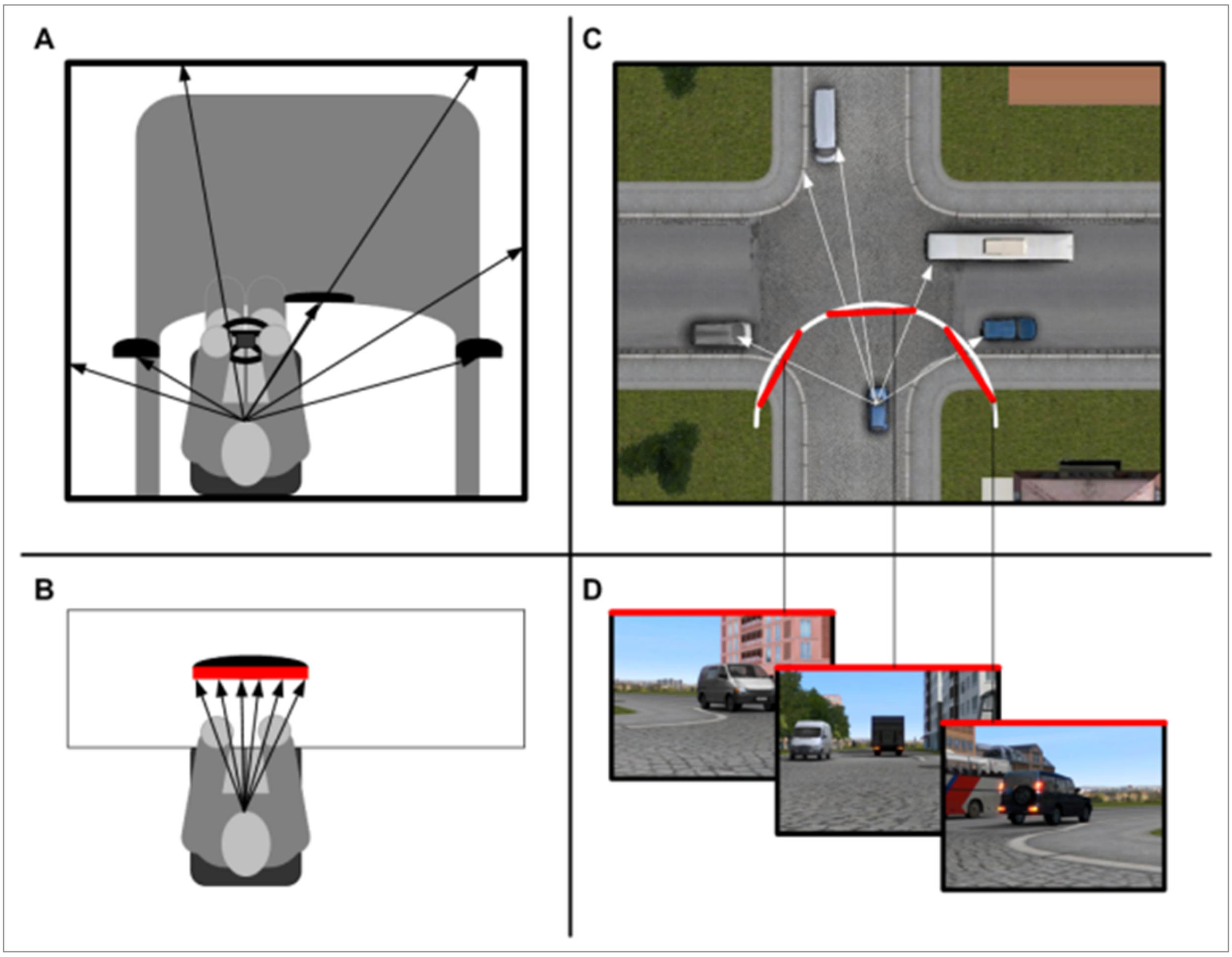

Fig. 16: Real visual environment (top) and possibilities for PC presentation (below)

Only a small section of this real visual environment can be reproduced on a computer screen without distortion. Realistic presentation of the driver's whole visual environment is thus not possible. Consequently, the driver perspective used on the screen cannot simply mimic the real driver perspective; instead, it is necessary to determine which of the available presentation options is best suited for the purpose in hand – namely the illustration of traffic situations for test questions. This limitation becomes evident from an analysis of the first variant developed, which is to be seen in Figure 17 (A).

The essential advantage of this variant is that the view appears relatively familiar at first sight. Closer consideration, however, reveals also numerous disadvantages. The first, for example, is the impression that the viewer is sitting not in the driver's seat, but in the centre of the back seat of the vehicle. Furthermore, this variant uses only approximately one-third of the available screen area of a conventional 4:3 monitor. The dashboard instruments and mirrors are also very small, which means that any relevant cues there are difficult to detect. The actual purpose of the illustration,

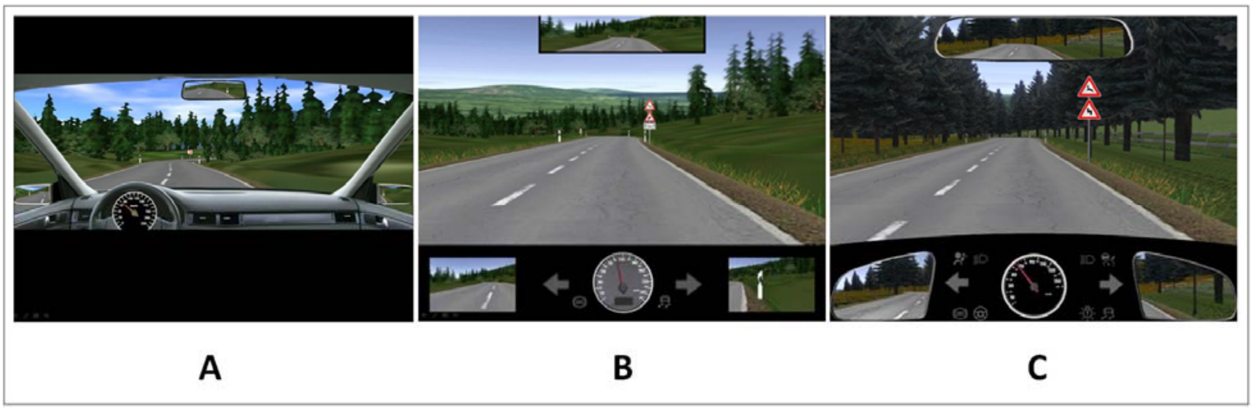

Fig. 17: Possibilities for dashboard instrument and mirror presentation as elements of an abstract vehicle interior

namely the clearest possible novice-oriented presentation of the cues relevant for test questions, thus cannot be satisfied optimally. The next step was therefore to search for alternatives which were more appropriate to the purpose of the illustration. The first outcome of this process was the presentation variant shown in Figure 17 (B). The mirrors and dashboard instruments are here depicted as elements of an abstract vehicle interior. This detracts slightly from the impression of sitting in a real vehicle. On the other hand, the candidate is in any case aware of sitting not in a vehicle, but at a PC. For this reason, greater importance was attached to the clearest possible presentation of the traffic situation, rather than to realistic reproduction of the vehicle cockpit. Moreover, the abstract depiction permits utilisation of the whole screen area, and the cues can also be realised as larger screen objects. In addition, the size and form of the mirrors and dashboard instruments can be varied and matched optimally to the purpose of the illustration. In the end, the variant with rounded mirrors to be seen in Figure 17 (C) was chosen; this variant was considered more pleasing to the eye by most of the experts. A further advantage offered by this more abstract reproduction of the driver's perspective is that it can be used in the test for different licence classes (requiring only the removal of the rear-view mirror for certain classes).

4.2.2 Elaboration of test questions with dynamic traffic scenarios

On 4th July 2007, the development and evaluation group set up a working group “Test item development”, in which representatives of the Technical Examination Centres, driving instructors and the scientific community were brought together to develop examples for new test item formats. After weighing up different question types and different instruction and answer formats, the working group concentrated initially on the development of questions of the existing multiple-choice type, but now with a computer-generated video sequence to illustrate the traffic situation addressed by the question instructions. Within this process, due consideration was given to a number of fundamental decisions already taken within the development and evaluation group with regard to the situation content (e.g. no dangerous behaviour of the ego-vehicle), the dynamic situation presentation (opening image at the beginning, sequence can be viewed three times, last image remains on the screen) and the prescribed sequence for processing of the elements (no longer possible to return to the film when answering the question). One of the first decisions of the working group related to the length of the dynamic scenarios. Depending on the complexity of the situation or the number of hazard cues to be accommodated, different lengths of time are necessary to present the entire traffic situation. Where the illustration sequences end with a multiple-choice-style question, however, the different lengths are more likely to irritate the candidate. This has been demonstrated with examples from learning media (CD-DRIVES, COCKERTON & ISLER, 2003), according to which disadvantageous discrepancies may arise between the actual sequence length and the length expected by the viewer. On the basis of experimental experience, the working group members therefore specified a common length of approx. 15 seconds for the dynamic presentation. By the end of 2010, 18 question examples had been developed in accordance with these decisions to illustrate the benefits of dynamic traffic scenarios for situation illustration.

A number of the elaborated examples are based on questions which have already been used in the TDT in the past, but were identified as being in need of revision within the framework of the ontinuous evaluation (cf. Chapter 3). Others were derived from accident analyses or proposals relating to necessary test contents from working group members. In the course of the development work, the narrow limitations for the assessment of further competences such as hazard recognition also became evident; these deficits cannot be overcome within the framework of a multiple-choice question. The examples were nevertheless able to demonstrate the benefits of dynamic illustration with regard to the self-

Fig. 18: Questions elaborated by the responsible working group to date

explanatory nature of instructions (e.g. for the question addressing “zip-merging”) and the presentation of implicit hazard cues (e.g. the question with playing children), which in the opinion of the working group supplies convincing justification for the introduction of the new instruction formats.

4.3 Studies and measures relating to the elaboration and use of computer-generated illustrations

To provide a scientific foundation for the implementation of new instruction formats, and to answer questions regarding the manner of this implementation, various empirical studies were undertaken and are in some cases still ongoing. In the context of the present report, these studies are to be described only briefly in the form of an overview. Those interested in more detailed information are referred to the corresponding research reports, which can be made available by the TÜV | DEKRA arge tp 21 working group upon request.

4.3.1 Study on the recognition of variants of static illustrations

One of the benefits of computer-generated (static) images described in Section 4.1 lies in their efficient variability. Targeted variation of the depicted road users and traffic environment reduces the recognisability of the image and thus hinders answering of the question merely on the basis of memorised features in the image. The goal must be to force the TDT candidate to actively analyse the depicted traffic situation, such that the question can only be answered through the purposeful application of general rules and experience to the individual case.

Nevertheless, no empirically founded knowledge was yet available for the specific application described here to determine the image elements which must be varied and the extent of variation necessary to reduce recognisability without altering the perceived content of the image.

The first results of a study conducted by the TÜV | DEKRA arge tp 21 working group show that minor changes to the traffic environment do not hamper recognition of an illustration. It is rather the case that extensive modification of the traffic environment is necessary. This can still be achieved without affecting the comparability of the traffic situation, however, as comparability is guaranteed by the unchanged relative positions of the traffic participants, for example.

A follow-up study is to supply more detailed information on the particular impacts on recognisability from the variation of individual details of a pictorial image. As this study is not yet completed, it is for the moment only possible to refer to the still unpublished research report “Recognition of varied static depictions of traffic situations” (FRIEDEL, 2010).

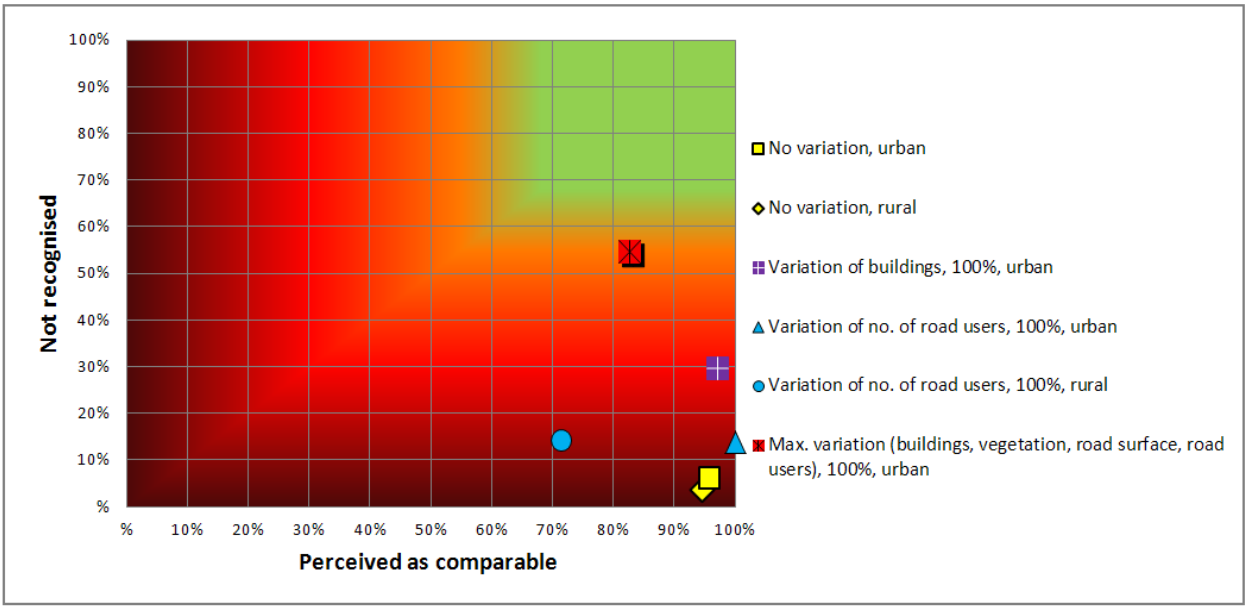

Fig. 19: Impact of variation on recognition and content comparability

4.3.2 Scientific expertises on new test contents

The “System Manual on Driver Licensing (Theory Test)” specifies that, prior to the elaboration of draft questions relating to test contents which have not been included or taken into account in methodically acceptable form in the question catalogue to date, a scientific expertise is to be produced on the content-related and methodical background to the question design. As the working group responsible for test development had elaborated a question on “Observing motorcyclists when turning”, a topic which is not included in the current question catalogue, this scenario became the subject of such a scientific expertise. The expertise explains the significance of this aspect for novice driver preparation, provides reasoning for the inclusion of this test content into the TDT, and gives recommendations for methodical implementation of the question; these points are sketched briefly in the following.Motorcyclists represent a group of traffic participants subject to an especially high risk of injury or death. The car driver is often seen as the principally responsible party in motorcycle accidents involving a car. Possible reasons here are deficits in hazard perception and judgement, which can be derived from the particular mechanisms of perceptual psychology effective in the interactions between car drivers and motorcyclists. Given the high level of risk for motorcyclists which emanates from car drivers, and considering the generally recognised scientific recommendation that a significant contribution to a reduced accident risk for motorcyclists could also be achieved with the training and testing of car drivers, it can be ascertained that hazard perception and judgement in connection with motorcyclists has not yet been taken into account adequately in the official catalogue of test questions. It was thus recommended that the official catalogue be expanded by elaborating corresponding questions, for which instruction formats with computer-generated dynamic scenarios will probably be used. Typical traffic situations which accident analyses have identified as particularly difficult for adequate hazard perception and judgement in connection with motorcyclists are to serve as the basis for question design.

4.3.3 Study of user satisfaction with dynamic driving scenarios

Dynamic driving scenarios in the form of computer-generated film sequences are in future to help improve the assessment of those driver competences which are essential for safe participation in road traffic (e.g. with regard to hazard avoidance). First of all, corresponding innovative question prototypes were developed on the basis of the conventional multiple-choice question format. Before commencing experimental studies to evaluate the quality of these question prototypes from the perspective of psychological testing, it is necessary to gather information regarding their acceptance on the part of the user, i.e. the driving licence applicant. An empirical study (FROMMANN & GENSCHOW, 2009) was thus conducted to assess user satisfaction with dynamic driving scenarios. The objective was to determine whether and to what extent questions with dynamic driving scenarios are able to satisfy the intention of a realistic presentation of the test contents in a form readily understandable for the test candidate.

As hardly any data was available on user satisfaction with dynamic driving scenarios, an explorative methodical approach was chosen for the study: In a series of partially standardised interviews, a sample of 20 test candidates was asked to evaluate a particular dynamic driving scenario (“zip-merging”).

The survey instrument developed for the study was based on the standardised design principles for ergonomic multimedia software user interfaces (DIN EN ISO 14915-1, 2002). The study showed that, from the perspective of most study participants, the presented dynamic driving scenario “Zip-merging” was very much in line with the standardised criteria for ergonomic software interface design. Almost all the interviewees described the dynamic presentation as “very realistic”, “very clear and understandable” and “very motivating” (to view further such traffic scenarios). The majority of the interviewees were of the opinion that the presented traffic scenario was suitable as a basis for answering of the test question. More detailed results can be found in the report “Study of user satisfaction with dynamic traffic scenarios” (FROMMANN & GENSCHOW, 2009).

Fig. 20: Test question with dynamic driving scenario (“zip-merging”)

The results of the explorative study permit only conditional generalisation of the findings. Firstly, exclusively relatively young driving licence applicants of ages under 30 years were questioned – it cannot be said whether and to what extent the dynamic driving scenarios meet the demands placed on software interface design from the perspective of older test candidates, or whether specific design deficiencies are present with regard to this user group. Secondly, the study referred only to one dynamic driving scenario. Even if it can be derived from the explorative study that there are apparently no design deficiencies in respect of certain universal features which are used in identical fashion in different question prototypes (e.g. the dashboard instruments of the ego-vehicle), possible deficiencies in the question-specific design features of the different question prototypes must be investigated separately. Consequently, a usability study should always be conducted for all question prototypes with dynamic driving scenarios.

4.3.4 Survey among driving instructors on questions with dynamic situation presentation

Alongside the positive subjective assessments of the questions with dynamic situation presentation by driving licence applicants (cf. Sections 4.3.2, 4.3.5 and 4.3.6), the opinions of driving instructors were also surveyed, not least in view of the impact which introduction of this instruction format will have on corresponding areas of professional driver training (control function of the test).

The Federation of Driving Instructor Associations (BVF) permitted the TÜV | DEKRA arge tp 21 working group to have four newly developed questions with dynamic situation presentation assessed by way of a questionnaire survey among driving instructors attending the 3rd German Driving Instructor Congress on 26th/27th November 2010. The survey respondents were able to give detailed feedback on the dynamic presentation of the traffic situation and on the associated instructions and answer options. To summarise, the results show that each of the evaluated questions was considered “suitable” by at least 96 per cent of the 53 driving instructors asked. Two questions were even considered “suitable” without exception. The questions concerned related to situations where a hazard cue was initially visible, but was then concealed shortly before a possible collision (e.g. children hidden behind parked cars). A detailed presentation of the methodical approach and results is to be found in the report “Survey on test questions with dynamic situation presentation among driving instructors attending the 3rd German Driving Instructor Congress” (BRESSENSDORF, FRIEDEL, GLOWALLA, RÜDEL, TSCHÖPE, WAGNER & WEIßE, 2011).

4.3.5 Study of situation-specific benefits of dynamic presentation

The benefits of dynamic presentations of traffic situations mentioned in Section 4.1 cannot be assumed for all traffic situations to be depicted for purposes of the TDT. There are some questions for which visualisation by way of a static image is still sufficient. Nevertheless, there are certain cases where the experts are disagreed as to whether a dynamic presentation of the traffic situation offers clear benefits compared to a static image, or whether the latter is adequate.

To be able to base future discussion on empirical data, the Chair of Traffic Psychology at the Dresden University of Technology is conducting a corresponding study. The starting point is the question as to the traffic situations in which a dynamic presentation is advantageous compared to a static image. The study provides for questions with static images and questions with dynamic presentations to be assessed by experts and novices (driving school students); their corresponding eye movement parameters are recorded and evaluated accordingly.

Initial results indicate that novices tend to find it easier to identify the traffic-relevant information (e.g. hazard cues) in a dynamic presentation than in a static image. As far as situation and hazard anticipation is concerned, the description of the further development of a situation appeared much more straightforward for the driving school students in the case of a dynamic presentation. Moreover, 93 per cent of the study participants (experts and novices) viewed the potential use of dynamic presentation in the TDT as “good” or “very good”.

A final answer regarding those traffic situations which would benefit particularly from dynamic presentation and those for which a static image appears adequate cannot be given until all the data have been evaluated – probably in late 2011.

Fig. 21: Examples for the eye movement of novice drivers when viewing test questions with dynamic driving scenarios

4.3.6 Eye tracking with dynamic situation presentations

Within the framework of the development of innovative, dynamic question formats, the design of the virtual cockpit, including mirrors (cf. Section 4.2), was the subject of an empirical study. The intention was to determine whether the simultaneous presentation of the forward and rearward traffic situation (the latter via rear-view mirrors) on a computer monitor leads to distraction or non-targeted traffic observation. The study “Characterisation of the eye movements of inexperienced novice drivers” (KORLUß, 2010), which was completed at the Chair of Traffic Psychology at the Dresden University of Technology in May 2010, analysed the eye movement of driving school students by tracking their point of gaze when viewing the questions with dynamic driving scenarios developed so far.

The results show that the dashboard instruments and mirrors were observed in a manner appropriate to the depicted situation. That means, for example, that in scenarios in which the traffic behind the vehicle was immediately relevant for the given situation, the mirrors were more frequently the target of the viewer's gaze than in scenarios with a different thematic focus. The same applies for the dashboard instruments. The possibility to view the scenarios repeatedly (here three times), in particular, resulted in increasingly appropriate eye movement. In all the evaluated scenarios, the viewer concentrated primarily on the forward traffic situation, interrupted by brief fixation of the dashboard instruments or mirrors. The results give reason to assume that the novice drivers are not distracted by the depiction of the mirrors, and that they tend to use them in accordance with the current situation – as would be the case in real motorised road traffic.

4.3.7 User testing after completion of the regular TDT

The objective for the testing of innovative formats within the framework of the regular TDT is to assess the methodical quality of the newly developed questions (formatively) already during the development process. On this basis, the questions can then be revised, if necessary. Testing thus effectively represents an advance evaluation of the questions before their potential official approval for use in the TDT.

The first step was to explain the objectives and significance of the testing to the responsible transport ministry in Saxony, which resulting in ministerial approval for the study. In 2009, DEKRA agreed to support the study at its Dresden branch office. Subsequently, the Saxon Association of Driving Instructors was informed about the project and similarly asked for its support. The regional driving schools concerned then received information on the objective and procedure for question testing.

Once the necessary software had been integrated into the existing DEKRA EDP system and successfully tested, the study was conducted in March 2010. The methodical approach is described in detail in the report “Testing of innovative questions within the framework of the regular TDT – Methodology” (FRIEDEL, WEIßE & RÜDEL, 2010). In the following, this approach is here outlined in brief.

First, each candidate completed the official test. After the candidate had submitted the completed test sheet, a short text was displayed to inform about the study. If the candidate then agreed to participate voluntarily in the study, three newly developed questions were to be answered. The three questions were presented within the familiar design framework of the official test. After answering, the candidate received feedback on the additional questions and was asked to provide a few anonymous personal data. Subsequently, the official test was evaluated by the official examiner. Study participation remained anonymous and had no influence on the completion, performance assessment and test decision of the regular TDT. Questions with dynamic situation presentation were tested in conjunction with a total of 523 official theoretical tests. The anonymised data records were passed on to the TÜV | DEKRA arge tp 21 working group for evaluation. On the basis of these data, question-specific profiles were produced for use for a first methodical assessment of the question concerned.

The results cannot be generalised and are thus illustrated here by way of three examples. A question relating to zip-merging, for instance, was answered correctly by 95 per cent of the candidates, compared to only 85 per cent who provided a correct answer to a question from the regular question catalogue with comparable content but purely textual descriptions. Assuming content equivalence for the two questions, this finding indicates that a dynamic presentation of the traffic situation greatly simplifies its interpretation. A second question with a dynamic scenario relating to the hazard of playing children was answered correctly by 77 per cent of the candidates for a first driving licence for vehicle class B, whereas the current official question with an image containing solution hints was answered correctly by all the candidates. This suggests that the intended assessment of hazard perception competences is better achieved by the new question with no solution hints remaining visible in the final image. This assumption is supported by the fact that the probability of a correct solution to the new question increases to 90 per cent among the candidates for higher licence classes with corresponding driving experience. A question relating to consideration towards weaker road users (pedestrians and cyclists) when turning right, on the other hand, received a correct answer from only 51 per cent of the candidates for a first licence for vehicle class B; among the candidates for a higher licence class, the question was answered correctly by 70 per cent. For these questions, the authors must weigh up whether the low rate of correct answers is a result of an ambiguous situation depiction which must thus be revised. The use of computer-generated scenarios makes it possible to introduce corresponding modifications into the depicted scenario. As these examples show, it is not possible to make general statements on the quality of all questions with dynamic scenarios; only question-specific statements can be derived.

As part of the question testing study, anonymised data were also collected on topics such as acceptance of the new question format. The data available to date indicate that around 71 per cent of the driving school students consider the questions with dynamic scenarios to be “better” or “much better” than the current question format.

4.3.8 Study on the recognisability of variants of dynamic scenarios

Compared to a static image, dynamic presentations contain more hazard cues (objects, sequences, etc.) by which they can be recognised. There is thus an even greater risk (cf. Section 4.3.1) that the correct answer in a particular scenario could be recalled from memory on the basis of such superficial cues, without requiring active analysis of the depicted situation (e.g. “That is the film where the correct answer is ‘Children’.”).

One possible solution to this problem − as similarly in the case of static images − lies in diverse variation of the dynamic situation. Alongside the possibility of varying the traffic environment chosen to depict identical content, it is also feasible to vary the content of the question while leaving the traffic environment unchanged. This would produce questions in which different situations are presented within one and the same environment, and in which different answers become appropriate in each case (“What must you pay attention to here?” Possible answers: “Children”, “Motorcyclist”, “Car”). The candidate is thus required to think carefully about the situation and must understand its meaning in order to be able to answer the question correctly.

A first study on this topic (“Changes in psychometric parameters in case of repetition and as dependent on the variation of questions”; HANDRIK, 2010) aimed to determine how the psychometric parameters applicable to a question are changed when that question is repeated, with and without variation of the question presentation. In the course of this study, questions with dynamic scenarios were answered repeatedly by groups of driving school students. One group answered exclusively different questions, while a second group was presented only a few different questions with constant variation of the content and depicted environment. The results show that the questions with dynamic scenarios are answered correctly by all the candidates after only a few repetitions. There was no significant difference between those who had learned with wholly different questions and those with only a few, but at the same time varied questions. This permits the conclusion that variations are not recognised more quickly than completely different questions. A broader study in 2011 is intended to provide more reliably founded data.

4.4 Conclusions from the studies on computer-generated instruction formats

If image variants are used to prevent the schematic memorisation of questions based on static images, the fullest scope of possible variation should be exploited, as minor changes (e.g. switching of vehicle colours) have practically no impact on the recognisability of the illustrations. Computer-generated illustrations, however, display greater similarities than the photographs used to date due to their uniform perspectives and a single library of objects (ground layouts, vehicles and building objects) for use across all illustrations. Consequently, greater interference can be expected with regard to perception of the different illustrations, placing an additional obstacle in the way of rote learning.

On the basis of the present study results, it can be assumed that the instruction formats developed for multiple-choice questions with dynamic situation presentation are readily understandable for the test candidates and also meet with their broad acceptance. The same applies for the driving instructors questioned. Novice drivers tend to identify traffic-relevant information better in the case of a dynamic presentation, and consequently consider the situation description to be more understandable compared to text or photographic illustrations. This finding was also confirmed by eye tracking analyses, which showed that, when viewing a film sequence, the candidates concentrated their attention on those objects which the question authors also deemed relevant for interpretation of the traffic situation. The quality of individual questions, however, can only be evaluated fully through testing subsequent to the actual theoretical driving test, which should thus always serve as the basis for submission of a question for approval.

The problem of schematic memorisation is as relevant for instruction formats with dynamic scenarios as it is for the use of static images. It can even be assumed that, due to the additional information provided by the development of a situation over time, it is generally easier to memorise a dynamic situation. This effect can be countered by defining as many corresponding questions as possible. The scope of such a catalogue of questions, however, remains limited by the number of suitable situations and the required test contents. Nevertheless, studies of the recognisability of dynamic scenarios indicate that multiple combinations of traffic situations and traffic environments are able to produce corresponding interference. This hinders the rote learning of scenario/solution pairs, and in this way makes a sound contribution to reducing the likelihood of schematic memorisation.

4.5 Steps to realise the introduction of computer-generated instruction formats

On the basis of the described findings, a resolution was passed at Meeting II/2010 of the Federal/Regional Expert Committee “Driver Licensing and Driving Instructor Legislation” (BLFA-FE/FL) in Schwerin on 29th/30th September 2010, providing for all illustrations of traffic situations in the corresponding questions of the official catalogue to be replaced with new, computer-generated images with effect from 1st July 2011. The TÜV | DEKRA arge tp 21 working group had previously developed corresponding computer-generated versions of all the illustrations of traffic situations used in the official catalogue of questions. Pursuant to the BLFA-FE/FL resolution, these images were published as amendments to the question catalogue for the TDT at the end of 2010, namely in issue 24/2010 of the official transport ministry gazette (“Verkehrsblatt”).

At Meeting I/2011 of the BLFA-FE/FL in Celle on 23rd March 2011, the latest knowledge relating to test items with dynamic instruction formats was presented. Subsequently, a policy decision was taken to introduce corresponding items from 2012.

In preparation of this step, the working group for test item development continues to focus on the elaboration of an adequate pool of questions. At the same time, scientific evaluations are being produced for all potential questions, and the questions themselves are being tested within the framework of the TDT. To this end, the corresponding testing has been extended to involve all Technical Examination Centres. The results are summarised in enclosures to the submissions for official approval, in accordance with the methods and procedures specified in the “System Manual on Driver Licensing (Theory Test)”. The databases and software systems to support the voting and approval procedures have also been adapted accordingly to permit the displaying of test items based on dynamic scenarios.

Mathias Rüdel, Dietmar Sturzbecher, Jan Genschow & Bernd Weiße

5 Development of innovative test items and forms of testing for extended driving competence assessment

It was already pointed out in the opening chapter – on the basis of details from the "System Manual on Driver Licensing (Theory Test)" – that innovation reports serve to present the results of studies and the research and development work aimed at the optimisation of test item formats and forms of testing. Furthermore, information is to be provided on the planned further development of the overall system of driving licence testing. Following up the presentation of essential results obtained so far in the two previous chapters, namely the findings from methodical work on optimisation of the present TDT done by the TÜV | DEKRA arge tp 21 working group, the focus is here shifted to planning for the development of innovative test items and forms of testing, alongside certain additional information on continuation of the aforementioned optimisation work. This involves questions as to the possible design of optimised models for novice driver preparation and of improved procedures and forms of testing, as well as determination of the prerequisites to be established for their implementation.

Information on continued optimisation of the present TDT

The introduction of computer-generated static images as situation illustrations for question instructions, which was decided at the end of 2010 and has in the meantime been realised, represents an important milestone: Prototypical "original" illustrations of traffic situations are already available in different variants for a large number of test questions and serve to hinder the schematic memorisation of question contents. During the next report period, such variants are to be produced to cover the majority of the image-based question instructions. In addition, numerous examples for test questions with dynamic scenarios to illustrate the question instructions were elaborated during the report period. On the basis of these examples, the meeting of the BLFA-FE/FL committee in March 2011 reached a policy decision on their introduction from 2012. A further task for the coming report period is thus to build up the necessary pool of questions and to subject these questions to formative and summative evaluation. The restructuring of the parallel tests and simplification of the present weighting system for the assessment of test performance, as recommended on the basis of the scientific evaluation, are to be realised after the introduction of questions with dynamic instructions. These development steps require amendment of the corresponding regulations and guidelines governing driver licensing. Such amendments could be made in the course of implementation of the Third EU Directive on Driving Licences in 2013. Specific recommendations to this end are to be elaborated during the next report period.

Work and planning for the development of innovative test items and forms of testing

As already mentioned in Chapter 2, the Driving Licence Regulations (§ 16 FeV) expect the TDT to furnish proof that the licence applicant is acquainted with the hazards of participation in motorised road traffic and with the behaviour necessary to avoid such hazards. Long-standing expert opinions (HAMPEL, 1977), and not least the results of the revision project, indicate that further optimisation is still necessary to meet this demand of the regulation authors: The testing of content from the subject area "Hazards", and in particular the assessment of capabilities relating to traffic perception, hazard recognition and hazard avoidance, can and must be improved with the aid of computer-assisted, innovative test items and new forms of testing.

To explore the methodical design possibilities for computer-assisted test items and formats, the TÜV | DEKRA arge tp 21 working group elaborated a catalogue of different item types with different instruction and answer formats at the beginning of the report period and provided examples for some of these types. This catalogue was developed further as part of a CIECA project involving also experts from other test organisations, and documented in the report on the work package WP300 of the CIECA Theory Test Project (WEIßE et al., 2009). The illustration formats ranged from simple true/false questions to driving simulation, and included, among others, multiple-choice questions, questions requiring free inputs, situation comparisons and tasks serving to determine reaction times. All of these formats display methodical advantages and disadvantages, as similarly described in the aforementioned report. These advantages and disadvantages, however, were evaluated very differently by the experts at national and international level, also as a result of the different priorities assigned to the assessment criteria used. The criteria in question related to economic aspects, the integration of the test items into a particular system of novice driver preparation, legal frameworks or purely scientific considerations. Given these evaluation discrepancies – which are by all means understandable in light of the different historical backgrounds to the individual systems of novice driver preparation – it was decided to abandon the original objective of the CIECA project, namely the development of common concepts for future test item formats in the sense of "best practice". The continuous exchanges of experience and research results, however, are to be maintained. Consensus was also reached on the standpoint that the possibilities of computer-assisted designs should be exploited further in the future. This includes not least the targeted use of dynamic driving scenarios in countries such as Belgium, Luxembourg, Estonia, Latvia, the Czech Republic and Hungary, as a means to lend the test items a more realistic appearance and to permit more valid assessment of the test candidate's driving competence (WEIßE et al., 2009).

How is the theory test realised at international level? In their comparison of novice driver preparation in 44 countries, Genschow, Sturzbecher & Willmes-Lenz (in press) determined that the computer is the predominant test medium. With regard to content, however, the computer-assisted tests often resemble the earlier paper-and-pencil tests, i.e. the test sheets have merely been transferred to the computer. The test demands are generally limited to a verification of declarative knowledge, which is achieved by way of true/false and multiple-choice questions with textual instructions and static pictorial illustrations. The potential of the test medium computer for the development of innovative instruction and answer formats to assess implicit knowledge is currently utilised in only a few countries. Instruction formats with dynamic driving scenarios, as they are to be used in Germany from 2012, can at present be considered pioneering exceptions (e.g. animated turn indicators in France).

In (1) the United Kingdom, (2) the Netherlands and the Australian states (3) Queensland, (4) New South Wales and (5) Victoria, hazard perception tests have been developed as an independent innovative form of testing to assess the corresponding components of driving competence and other implicit knowledge related to the simulated traffic and driving situations. The methodical design of these hazard perception tests and their integration into the national systems of novice driver preparation are outlined in the following with excerpts from the study report of Genschow, Sturzbecher and Willmes-Lenz (in press).

1. In the United Kingdom, driving licence applicants must take a hazard perception test immediately after the theory test (i.e. before the practical driving test and before the commencement of solo driving). This test was already introduced in 2002 and comprises 14 one-minute video sequences, in which a drive is presented from the driver perspective. As soon as the candidate recognises a hazardous situation, this must be indicated by clicking with the mouse. Thirteen of the video sequences contain exactly one hazard cue. In one case, there is also a second hazard cue to be identified; the candidate, however, is not informed as to the film sequence concerned. Test performance is assessed by way of the candidate's speed of reaction. A maximum of five points is awarded for each hazard to be identified, depending on the speed of the corresponding reaction. The assessment is weighted according to how early the development of a hazardous situation is recognised and when the candidate reacts accordingly. With a total of 15 hazard cues to be identified, the maximum attainable score is 75 points; the candidate must collect at least 44 points to pass the hazard perception test.
2. In the Netherlands, novice drivers must take a computer-assisted test (during the supervised learning phase, before taking the practical driving test) which involves firstly a hazard perception test (Part 1) and immediately thereafter a traditional theory test (Part 2). To pass the first part of the test, at least 12 of the total of 25 questions must be answered correctly. These questions comprise hazard-related photographs depicting traffic situations from the driver perspective (with information in the mirrors and with turn indicators and speedometer visible), on the basis of which the candidate must select one of the following three reaction options: “Apply the brakes”, “Take foot off the accelerator” or “Do nothing”. The corresponding answer to each question must be given within a time of eight seconds. To pass the second part of the test, 35 of a total of 40 questions (true/false questions, multiple-choice questions, numerical inputs) must be answered correctly; between 8 and 15 seconds is allowed to answer each question, depending on the question format. The question instructions are in part illustrated by way of drawings or photographs, either from the driver perspective or as a bird's eye view. After completing both parts of the test, the novice driver receives corresponding feedback, for example on those questions of the hazard perception test which were answered incorrectly, too late or not at all.
3. In the Australian state of Queensland, successful completion of a hazard perception test is a prerequisite for the graduation from a “P1 licence” to a “P2 licence”; this graduation lifts certain protective regulations applicable to solo driving. The hazard perception test is an online test and is intended to measure the driver's ability to anticipate, recognise and react appropriately to different hazardous situations in road traffic. The hazards to be identified relate to possible collisions with other road users (e.g. other motor vehicles, pedestrians, cyclists) which the candidate could avoid by either slowing down or changing course. The test consists of a series of real-life film clips, in which the candidate must click on the screen with the mouse to indicate the particular feature of the traffic situation which demands a particular response by the driver. It records both whether a potential hazard is identified and how quickly the candidate reacts to the hazard cue.

4. In the Australian state of New South Wales, a hazard perception test is stipulated at the earliest 12 months and at the latest 24 months after the start of solo driving; if the test is passed, some of the special protective regulations applicable for novice drivers are lifted. A second hazard perception test must be taken 24 months later and leads to the granting of a driving licence no longer subject to protective regulations. The hazard perception tests are intended to assess whether the novice driver is able to recognise potentially hazardous situations and react appropriately. Three separate content-related demands are distinguished: Observance of the necessary safe distance to other vehicles, selection of appropriate safe gaps when negotiating junctions or changing lanes, and the identification of hazards in front of, behind and to the side of the driver's own vehicle. The traffic and hazard situations in the real-life film clips presented are based on the five most common types of accident involving novice drivers.

   The first hazard perception test comprises 15 questions with 30-second real-life videos, in which the novice driver must touch the screen as soon as he deems it safe to perform a certain action. Before each film sequence, an 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”). Each question involves one single action. Possible actions are “Slow down”, “Overtake” or “Turning/crossing at a junction”. The candidate is also able to see the speed of his own vehicle and the activation of the vehicle's turn indicators during the film clip. The actual test is preceded by two practice questions to enable to candidate to become familiar with the test procedure. At the start of each film clip, a still image is displayed for a few seconds. Once all 15 questions have been answered, a message appears on the screen to tell the candidate whether the test was passed. If the test is passed, the candidate also receives feedback on areas in which existing driving competence should be improved; if the test is failed, the feedback points to the areas in which further practice is necessary before attempting the test again.

   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). Unlike the first hazard perception test, the film clip does not necessarily contain only one instance of the required action; there may be several occasions on which the given action can be performed. Ten film clips are presented in total, each of which is longer than those in the first hazard perception test. This second hazard perception test is a component of the so-called “Driver Qualification Test”, which begins with a traditional knowledge test comprising 15 multiple-choice questions. The pass rate in the “Driver Qualification Test” is around 67 per cent.

5. In the Australian state of Victoria, novice drivers are required to pass a hazard perception test during the supervised learning phase, before being allowed to attempt the on-road practical test. The candidate is shown a total of 28 videos of traffic situations from the driver perspective. Before each video, the required driving action is specified (slow down, overtake, turn or move off). The candidate is asked to decide when this required action can be performed safely during the given driving scenario. Each clip begins with a verbal description of the coming situation (e.g. “You are driving straight ahead”), followed by a still opening image of the traffic situation from the driver perspective, in which the current speed of the vehicle is shown on the speedometer. The subsequent screen view points ahead to the action (see above) which will be relevant in the clip (e.g. “Click the mouse when you think you should slow down”). The still opening image then returns and starts to flash to indicate that the video is about to begin. During the test video, the novice driver must “perform” the specified action by clicking the mouse as soon as a suitable moment arises.

The above examples show that the hazard perception tests implemented in the individual countries differ significantly in terms of their methodical design, instruction formats (e.g. video sequences, virtual traffic scenarios, static images) and required response (e.g. mouse click in response to hazard cues, selection of an action decision); this also means that different demands are placed on the driving licence applicant. In all cases, however, these demands go far beyond those of a traditional knowledge test which addresses declarative knowledge: The hazard perception tests assess evidently implicit knowledge of necessary driving behaviour in near-realistic traffic situations; this requires elements of competence corresponding to those defined by the information processing model (DODGE, 1982) described in Chapter 2 to be applied to different traffic situations, namely the ability to identify safety-relevant hazard cues, to interpret traffic situations (e.g. targeted information searching within limited time), or to clarify, weigh up and select an appropriate action.

The methodical potential of hazard perception tests seems apparent: Compared to traditional knowledge tests, they strengthen the situation reference of the test items and achieve a greater content correspondence between the demands of the test and those of driving in actual traffic (e.g. with regard to the identification of hazards), without exposing the still inexperienced candidate to real hazards. Hazard perception tests could be used − as already demanded by Bönninger and Sturzbecher in 2005 – to simulate a multitude of conceivable hazard situations and to operationalise their solution in standardised test items. The responses to such test items could be evaluated objectively according to different parameters (e.g. speed, rule violations, safe distances); this would further increase the objectivity of the driving licence test. Furthermore, different driving scenarios or test items could be joined together in sequences to simulate journeys with different hazardous situations. It would be a simple matter to record these journeys as a basis for the self-evaluation of driving behaviour by the candidate or for discussion with the driving instructor or examiner. At the same time, finally, it would be possible to test seasonal demands on driving behaviour (e.g. driving in snow) which cannot be tested within the framework of a test drive in real traffic.

It is at present not possible to say for sure, which new types of test items and which forms of testing for the assessment of implicit knowledge will be used from which point in the future in Germany. It only seems certain that − together with the research and development work of the TÜV | DEKRA arge tp 21 working group − basic research in this direction is indispensable. One important example from the report period is the BASt project 82.326 “Test psychology and learning theory underpinning the test questions in theoretical driving tests, with special consideration of test question formats with image sequences”, in which the University of Saarland proposed innovative test item formats and developed corresponding examples. It was recommended, for example, that reaction-based items be introduced: After an instruction in the form “Press the space bar once as soon as you identify a need to reduce your speed”, dynamic driving scenarios could depict a complex traffic and hazard situation. Such driving scenarios are created with the aid of a special authoring software (VICOM) from the TÜV | DEKRA arge tp 21 working group. The action reference of such items can be increased over the further stages of format development, for example by enabling the candidate to control the speed of his (simulated) own vehicle. Long-term preparations for the implementation of such formats in the TDT, however, still require decisions on numerous issues besides those addressed by the findings of current basic research. This relates, for example, to the elaboration of criteria for the selection of traffic scenarios and for the assessment of reaction times; it must be clarified, not least, whether the present approval procedures are still suitable to deal with these formats. Continuation projects at the Dresden University of Technology and the University of Saarland are to be initiated during the coming report period to consider such topics.

Looking ahead, finally, it must also be asked whether the assessment of factual knowledge by way of multiple-choice questions, on the one hand, and the assessment of implicit and action knowledge by way of corresponding new test item formats, on the other hand, should continue to be arranged at the same point in the process of driving competence acquisition and novice driver preparation – i.e. as elements of one and same test − or whether they should preferably be planned at different points along the timescale – i.e. as separate tests. This question concerns further differentiation of the present TDT and, in view of its selection and control functions, the appropriate placement of a further developed test and the possibly ensuing forms of testing (knowledge test and traffic perception test) within the system of novice driver preparation (see Chapter 2).

Which factors must be taken into account when answering this question? The present TDT, like other traditional knowledge tests assesses essentially factual knowledge, which can be acquired and tested even without deeper understanding of application-specific correlations. Consequently, successful completion of the test is already possible without prior driver training or at the very beginning of novice driver preparation. It is only at this point, however, that such a knowledge test can meaningfully achieve its control function: The factual knowledge must already be gained in advance of practical driving experience; once experience has progressed, other forms of testing could be used to verify factual knowledge in connection with traffic situations − a specific knowledge test, with its known methodical weaknesses, would then no longer be necessary. On the other hand, longer learning processes and longer periods of driving experience on the part of the candidate would be prerequisites for additional, more detailed assessment of implicit application- and situation-related knowledge within the framework of the TDT; a correspondingly modified TDT, or an additional traffic perception test, therefore, could not be taken immediately after theory classes, and would instead need to be arranged later in the system of novice driver preparation. These circumstances are also reflected in many graduated licensing systems, in which an initial test of declarative knowledge (e.g. knowledge of traffic rules and road signs) must be passed before being allowed to commence practical driving experience; the driving licence applicant acquires this knowledge independently with the aid of a manual or other learning media. The assessment of application-specific knowledge (e.g. evaluation of the hazard potential of a situation, weighing up of action alternatives to avoid hazards, reaction times) is left until a further theory test or hazard perception test at a later stage, by which time extensive practical driving experience has been gained.

In Germany, too, traffic perception tests could presumably also serve as a bridge between the traditional theoretical test and the practical driving test in the nearer future. These three forms of testing, with their specific methodical and content-related opportunities and limitations, should in future be meshed within the system of novice driver preparation, following the laws of developmental psychology as they apply to driving competence acquisition, and integrated accordingly into the corresponding teaching and learning forms. Hazard recognition and hazard avoidance are topics which cannot be tested properly in a traditional knowledge test or otherwise at the beginning of novice driver preparation (aforementioned lack of situation awareness due to insufficient driving experience), and are likewise inappropriate as elements of a traditional practical test, as the demand situations in real traffic cannot be controlled at will by the examiner, and hazardous situations also cannot be brought about deliberately for safety reasons. Traffic perception tests with freely computer-generated and standardisable traffic situations, on the other hand, are an outstanding means to offer precisely this possibility.

It appears necessary to understand the individual – existing and future − forms of testing as components of a more comprehensive methodical concept for the assessment of driving and traffic competence. Within this concept, they should ideally complement each other with regard to their test content and the elements of competence assessed, whereby each would compensate the methodical deficits and limitations of the other forms. One such limitation of the practical driving test, for example, is the fact that certain desirable driving tasks cannot be assessed at a given test location due to the lack of a corresponding road layout; this detracts from the validity and equality of the practical test. In the more distant future, these limitations could be overcome with a PC- or simulator-based test of process knowledge and driving behaviour in simulated driving situations – subject to the elaboration of realistic demand situations. Furthermore, contrary to a practical test in real traffic, a simulated drive permits targeted variation of the driving tasks in respect of further road and traffic conditions (e.g. weather, traffic density, etc.).

Nevertheless, it must not be concluded from the described methodical benefits of traffic perception tests and other forms of testing with simulated demand situations that the practical driving test in real traffic is expendable. On the contrary: Compared to a simulation, which only ever depicts an extract of overall reality, real traffic demands that the driver's limited attention resources be utilised under time pressure to continuously and simultaneously observe and judge a much wider scope of situation aspects in terms of their significance for the further development of the situation and its hazard potential. Besides information acquisition and information processing in the narrower sense, regulation of the appropriateness of possible reactions must be accomplished under incomparably more complex conditions in real traffic. And last but not least, awareness of the real hazards and the consequences of a loss of control in real traffic creates specific test conditions which suggest that a practical driving test is indispensable as ecologically valid proof (i.e. the test demands correspond to the real demands) of the level of driving competence reached within the framework of novice driver preparation.

Situation-related test item formats and forms of testing to assess traffic perception, hazard recognition and hazard avoidance are also a source of new challenges. This refers firstly to safeguarding of the validity of the test items: The more complex the demand situations and instruction formats, and the more action-referenced the answer format design, the less apparent the candidate's deeper understanding of the situation becomes (e.g. Why did the candidate click on a child as “hazard-relevant” in the virtual driving scenario? Simply because he saw a child, or because that child actually displayed hazard-related behaviour?). It must not be forgotten at this juncture, however, that there is also a far from insignificant probability of guessing the correct solution to a multiple-choice question without corresponding competences. Secondly, the demands relating to feedback on test performance increase with the complexity and action reference of the test items: To properly exploit the special learning potential of innovative test items assessing traffic perception, it is not sufficient to merely signal an incorrect solution or response; it should rather be explained to the candidate, why the chosen solution or response was not correct and how the situation should in fact be solved. In this respect, traffic perception tests are similar to the practical driving test. Thirdly, finally, the demands relating to adequate teaching/learning forms and the provision of corresponding teaching/learning materials increase with the methodical particularities of the test media. In other words: The driving licence applicant must be given sufficient opportunity to practise with learning materials similar in content and design to those used in the test; beforehand, it is also necessary to furnish empirical proof of the safety relevance and trainability of the demanded driving competences.

To summarise: The introduction of a PC-based theoretical driving test was for the TÜV | DEKRA arge tp 21 working group an opportunity to implement scientifically founded methodical improvements, as well as to elaborate and test innovative approaches to driving test design. This includes the development of instruction formats with computer-assisted dynamic driving scenarios, with which the determined methodical weaknesses of conventional multiple-choice questions are to be compensated to a certain extent in the future. At the same time, studies have been undertaken and research and development work has been initiated to improve the assessment of traffic perception, hazard recognition and hazard avoidance in the medium term, with innovative test item formats and perhaps also with new forms of testing such as a traffic perception test. This work is to be continued during the coming report period. To ensure the success of these efforts, both the networking with basic research in progress at other scientific institutions (Dresden University of Technology, University of Potsdam, University of Saarland) and international exchanges are to be expanded. At national level, the experience and planning of the TÜV | DEKRA arge tp 21 working group are to be integrated into the BASt project “Framework concept for novice driver preparation”, which is currently elaborating long-term design strategies for the overall system of novice driver preparation in Germany.

6 Bibliography

ANDERSON, J.R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

ANDERSON, J.R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum Akademischer Verlag.

BAUMERT, J. (1993). Lernstrategien, motivationale Orientierung und Selbstwirksamkeitsüberzeugungen im Kontext schulischen Lernens. *Unterrichtswissenschaft, 21,* 327-354.

BERNOTAT, R. (1970). Anthropotechnik in der Fahrzeugführung. *Ergonomics, 13*, 353-377.

BÖNNINGER, J., KAMMLER, K. & STURZBECHER, D. (Hrsg.). (2009). *Die Geschichte der Fahrerlaubnisprüfung in Deutschland*. Dresden: TÜV | DEKRA arge tp 21.

BÖNNINGER, J. & STURZBECHER, D. (2005). *Optimierung der Fahrerlaubnisprüfung. Ein Reformvorschlag für die theoretische Fahrerlaubnisprüfung. Berichte der Bundesanstalt für Straßenwesen*. Bremerhaven: Wirtschaftsverlag NW.

BRESSENSDORF, G. von, FRIEDEL, T., GLOWALLA, P., RÜDEL, M., TSCHÖPE, P., WAGNER, W. & WEIßE, B. (2011). *Befragung von Fahrlehrern zu Aufgaben mit dynamischer Situationsdarstellung im Rahmen des 3. Deutschen Fahrlehrerkongresses*. Dresden: TÜV | DEKRA arge tp 21.

BUNDESRAT (2007). *840. Sitzung des Bundesrates am 20. Dezember 2007. Empfehlungen der Ausschüsse. Siebzehnte Verordnung zur Änderung der Gebührenordnung für Maßnahmen im Straßenverkehr*. Accessed 09.09.2011. URL: http://www.fahrerlaubnisrecht.de/Pdf-dateien/BR-Drucksachen/BR-Drs.07-815%20-%2009.11.2007%20-%2017.VO%20% C4nd%20Geb%FChren-Ordnung.pdf.

COCKERTON, C. & ISLER, R.B. (2003). *Developing real-life driving simulations for novice driver education, ED-MEDIA 2003*.

DIN EN ISO 14915-1. (2002*). Software-Ergonomie für Multimedia-Benutzungsschnittstellen. Teil 1: Gestaltungsgrundsätze und Rahmenbedingungen*. Berlin: Deutsches Institut für Normung e.V.

DODGE, K. A. (1982). Social information processing variables in the development of aggression and altruism in children. In C. Zahn-Waxler, M. Cummings & M. Radke-Yarrow (Eds.), *The development of aggression and altruism: Social and sociobiological origins*. New York: Cambridge University Press.

DONGES, E. (1982). Aspekte der Aktiven Sicherheit bei der Führung von Personenkraftwagen. *Automobil- Industrie 27*, 183–190.

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

ENDSLEY, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors 37(1)*, 32–64.

FRIEDEL, T. (2010). *Wiedererkennung variierter statischer Bilder von Verkehrssituationen*. Unpublished research paper. Dresden: TÜV | DEKRA arge tp 21.

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

FROMMANN, P. & GENSCHOW, J. (2009). *Untersuchung der Nutzerzufriedenheit mit den dynamischen Fahrszenarien* (Berichte zur Evaluation der Theoretischen Fahrerlaubnisprüfung). Oberkrämer: IPV.

GEDIGA, G. & HAMBORG, K.-C. (1999). IsoMetrics. Ein Verfahren zur Evaluation von Software nach ISO 9241-1. In H. Holling & G. Gediga (Hrsg.), *Evaluationsforschung* (pp. 195–234). Göttingen: Hogrefe.

GENSCHOW, J., STURZBECHER, D. & WILLMES-LENZ (in press). *Entwurf des Forschungsberichts zum Projekt “Fahranfängervorbereitung im internationalen Vergleich”* (Vorläufiger Ergebnisbericht zum BASt-Forschungsprojekt FP 82.325/2007). Unpublished manuscript. Vehlefanz: IFK.

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”, M 201). Bremerhaven: Wirtschaftsverlag NW.

GREENO, J.G., COLLINS, A.M. & RESNICK, L.B. (1996). Cognition and learning. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.

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.

HAMPEL, B. (1977). *Einsatzmöglichkeiten audio-visueller Hilfsmittel im Rahmen der Führerscheinprüfung*. Bericht zum FA 7408 der BASt. Köln.

HAMPEL, B., JANITSCHKE, R. & SCHAFFRAN, K.H. (1977). *Entwicklung und Erprobung eines audio-visuellen Prüfungsinstrumentes für die theoretische Fahrerlaubnisprüfung*. Köln: TÜV Rheinland.

HANDRIK, M. (2010). *Verbesserung der theoretischen Fahrerlaubnisprüfung durch Lernen mit neuen Medien*. Unpublished research paper. Hamburg: Helmut Schmidt Universität – Universität der Bundeswehr

HATAKKA, M., KESKINEN, E., GREGERSEN, N.P. & GLAD, A. (1999). Theories and aims of educational and training measures. In: S. Siegrist (Ed.), *Driver training testing and licensing - towards theory based management of young drivers injury risk in road traffic* (Results of EU project GADGET, Work Package 3, BFU-Report 40, pp. 13-48). Bern: Schweizerische Beratungsstelle für Unfallverhütung.

HOFFMANN, L., BORNING, J. & STURZBECHER, D. (2009). *Überprüfung der softwareergonomischen Qualität des Bewerberprüfprogramms* (Berichte zur Evaluation der Theoretischen Fahrerlaubnisprüfung). Oberkrämer: IPV.

HORSWILL, M. & MCKENNA, F. (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*. Aldershot: Ashgate Publishing, Ltd.

HURRELMANN, K. (1999). Sozialisation. In G. Reingold, G. Pollak & H. Heim (Hrsg.), *Sozialisationsforschung* (pp. 481-486). München: Oldenbourg.

JOOS, M., RÖTTING, M. & VELICHKOVSKY, B.M. (2003). Die Bewegungen des menschlichen Auges: Fakten, Methoden, innovative Anwendungen. In G. Rickheit, T. Herrmann & W. Deutsch (Hrsg.), *Psycholinguistik/Psycholinguistics. Ein internationales Handbuch / An International Handbook*. Berlin & NY: de Gruyter.

KESKINEN, E. (1996). Warum ist die Unfallrate junger Fahrer und Fahrerinnen höher? In Bundesanstalt für Straßenwesen (Hrsg.), *Junge Fahrer und Fahrerinnen. Berichte der Bundesanstalt für Straßenwesen, Heft M52*. Bremerhaven: Wirtschaftsverlag NW.

KLIEME, E. (2004). *Der Beitrag von Bildungsstandards zur Qualitätssicherung und Qualitätsentwicklung in Schulen: Implementation, Weiterentwicklung und Nutzung der Standards*. Accessed 09.09.2011. URL: http://www2.dipf.de/publikationen/volltexte/klieme_ kmk_042004.pdf.

KLIEME, E. et al. (2007). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise* (unveränderte Auflage). Bonn/Berlin: Bundesministerium für Bildung und Forschung. Accessed 23.08.2011. 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 23.08.2011. URL: http://kompetenzmodelle.dipf.de/pdf/rahmenantrag.

KLUCKHOHN, C. (1951). Value and value orientations in the theory of action: An exploration in definition and classification. In T. Parsons & E.A. Shils (Eds.). *Toward a general theory of action* (pp. 388-433). Cambridge, MA.

KOEPPEN, K., HARTIG, J., KLIEME, E. & LEUTNER, D. (2008). Current issues in competence modeling and assessment. *Journal of Psychology, 216 (2)*, 61-73.

KORLUß, U. (2010). *Blickbewegungsmessungen.* Dresden: Technische Universität Dresden, Institut für Verkehrsplanung und Straßenverkehr.

LAMNEK, S. (2001). *Theorien abweichenden Verhaltens* (7. Auflage). München: Wilhelm Fink Verlag.

LEUTNER, D., BRÜNKEN, R. & WILLMES-LENZ, G. (2009). Fahren Lernen und Fahrausbildung. In H.P. Krüger (Hrsg.), *Anwendungsfelder der Verkehrspsychologie* (Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete, Serie VI Verkehrspsychologie, Vol. 2, pp. 1-79). Göttingen: Hogrefe.

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

MALONE, S., BIERMANN, A., BUCH, S. & BRÜNKEN, R. (2011) *Dynamische Darstellungs-und Antwortformate in der theoretischen Fahrerlaubnisprüfung*. Präsentation zum Abschlussworkshop des BASt-Projekts 82.326 "Testpsychologische und lehr-lerntheoretische Grundlagen von Prüffragen in der theoretischen Fahrerlaubnisprüfung unter besonderer Berücksichtigung von Prüffragenformaten mit Bildsequenzen" am 27.05.2011 in Saarbrücken, Universität des Saarlandes.

MAYCOCK, G. & FORSYTH, E. (1997). *Cohort study of learn and novice drivers. Part 4. Novice driver accident in relation to methods of learning to drive, performance in the driving test and self assessed driving ability and behaviour* (TRL Report 275). Crowthorne: Transport Research Laboratory.

McKNIGHT, A.J., & ADAMS, B.B. (1970). *Driver Education Task Analysis, Vol. I: Task Descriptions*. Alexandria, VA: Human Resources Research Organization.

McKNIGHT, A.J. & ADAMS, B.B. (1972). *Driver Education Task Analysis. Volume II: Task Analysis Methods. Final Report*. Alexandria, VA: Human Resources Research Organization.

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

OECD (2010). *PISA 2009 – Ergebnisse*. Accessed 09.09.2011. URL: http://www.oecd.org/document/53/0,3746,de_34968570_39907066_43433717_1_1_1_1,00.html.

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, 13 (3)*, 257-266.

ROTH, H. (1971). *Pädagogische Anthropologie Band II. Entwicklung und Erziehung*. Hannover: Hermann Schroedel Verlag.

SCHECKER, H. & PARCHMANN, I. (2006). Modellierung naturwissenschaftlicher Kompetenz*. Zeitschrift für Didaktik der Naturwissenschaften, 12 (1)*, 45-66.

SCHWENKMEZGER, P. & HANK, P. (1993). Papier-Bleistift- versus computerunterstützte Darbietung von State-Trait-Fragebogen: eine Äquivalenzuntersuchung. *Diagnostica, 39 (3),* 189-210.

SCOTT, W.R. (1995). *Institutions and Organizations*. Thousand Oaks, CA: Sage.

SHUELL, T.J. (1990). Phases of meaningful learning. *Review of Educational Research, 60 (4),* 531-547.

STURZBECHER, D. (2008). Methodische Grundlagen der praktischen Fahrerlaubnisprüfung. In D. Sturzbecher, J. Bönninger & M. Rüdel (Hrsg.), *Optimierung der praktischen Fahrerlaubnisprüfung. Methodische Grundlagen und Möglichkeiten der Weiterentwicklung*. Dresden: TÜV| Dekra arge tp 21.

STURZBECHER, D. (2010). Methodische Grundlagen der praktischen Fahrerlaubnisprüfung. In D. Sturzbecher, J. Bönninger & M. Rüdel, *Praktische Fahrerlaubnisprüfung – Grundlage und Optimierungsmöglichkeiten*. *Berichte der Bundesanstalt für Straßenwesen* (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. (2010). *Praktische Fahrerlaubnisprüfung - Grundlagen und Optimierungsmöglichkeiten. Berichte der Bundesanstalt für Straßenwesen.* Bremerhaven: Wirtschaftverlag NW.

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., KASPER, D. & GENSCHOW, J. (2009). *Einfluss der Randomisierung von Prüfungsaufgaben und Auswahlantworten auf die Prüfungsleistung* (Berichte zur Evaluation der Theoretischen Fahrerlaubnisprüfung). Oberkrämer: IPV.

STURZBECHER, D., MÖRL, S. & KALTENBAEK, J. (in press). Entwurf des Forschungsberichts zum Projekt „Optimierung der Praktischen Fahrerlaubnisprüfung" (Vorläufiger Ergebnisbericht zum BASt-Forschungsprojekt FP 82.345/2008). Unpublished manuscript. Vehlefanz: IFK.

TÜV | DEKRA arge tp 21 (2008). *Handbuch zum Fahrerlaubnisprüfungssystem (Theorie)*. Dresden: TÜV | DEKRA arge tp 21.

TÜV | DEKRA arge tp 21 (2011). *Referenzdokument zur Optimierung der Theoretischen Fahrerlaubnisprüfung.* Dresden: TÜV | DEKRA arge tp 21.

VAN Der MOLEN, H.H. & BÖTTICHER, A.M.T. (1988). A hierarchical risk-model for traffic participants. *Ergonomics, 31*, 537-555.

WEINERT, F.E. (1998). Neue Unterrichtskonzepte zwischen gesellschaftlichen Notwendigkeiten, pädagogischen Visionen und psychologischen Möglichkeiten. In Bayerisches Staatsministerium für Unterricht, Kultus, Wissenschaft und Kunst (Hrsg.), *Wissen und Werte für die Welt von morgen* (Dokumentation Bildungskongress, pp. 101-125). München.

WEINERT, F.E. (1999). *Konzepte der Kompetenz. Gutachten zum OECD Projekt „Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)"*. Neuchâtel, Schweiz: Bundesamt für Statistik.

WEINERT, F.E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (pp. 17–31). Weinheim und Basel: Beltz Verlag.

WEIßE, B. (2011). *Unfalldatenauswertung zur Optimierung der Theoretischen Fahrerlaubnisprüfung – Entwurfsfassung*. Unpublished research paper. Dresden: TÜV | DEKRA arge tp 21.

WEIßE, B., BURCH, I., LEWIS, J., MATTSSON, H., MEYER, R., HEUVEL, R., HOTTI, M., SANDERS, N., HIMA, T. & WENDEL-HELL, P. (2009). *CIECA Theory Test Project – WP300-Report*. Brüssel: CIECA.

WILLMES-LENZ, G. (2002). *Verringerung des Fahranfängerrisikos durch fahrpraktische Vorerfahrung*. Referat auf dem 38. BDP-Kongress für Verkehrspsychologie vom 12.-14. September 2002, Universität Regensburg, Regensburg. Accessed 23.08.2011. URL: http://psydok.sulb.uni-saarland.de/volltexte/2006/752/pdf/willmes_01.pdf.