

Computerbasierte Messung von Teilaspekten der Fahrkompetenz

Besonderheiten des Expertiseerwerbs beim Autofahren

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultät III
der Universität des Saarlandes

vorgelegt von
Sarah Malone
aus Saarbrücken

Saarbrücken, 2012

Dekan:

Prof. Dr. Jochen Kubiniok

Berichterstatter:

Prof. Dr. Roland Brünken

Prof. Dr. Susanne Buch

Prof. Dr. Detlev Leutner

Tag der Disputation: 5. Oktober 2012

Zusammenfassung

Eine vielversprechende Maßnahme dem deutlich erhöhten Verkehrsunfallrisiko, das Fahranfänger aufweisen (vgl. z.B. Maycock, Lockwood & Lester, 1991; Schade, 2001), entgegenzuwirken, besteht in einer Verbesserung der theoretischen Fahrerlaubnisprüfung. Insbesondere die Optimierung ihrer Validität erscheint relevant, da diese mit einer verbesserten Identifikation derjenigen Personen, die noch sicherheitsrelevante Defizite aufweisen, einhergeht. Das Ziel der vorliegenden Arbeit bestand darin, auf einer theoretisch und empirisch fundierten Grundlage Aufgabenformate zu entwickeln, die eine valide Trennung von Könnern und Nicht-Könnern ermöglichen. An solche validen Aufgaben werden die Anforderungen gestellt, dass sie zum einen sicherheitsrelevante Inhalte abprüfen und zum anderen eine adäquate Operationalisierung der gewählten Merkmale und Fähigkeiten darstellen. Die *Gefahrenwahrnehmung* hat sich in der Verkehrssicherheitsforschung als ein Aspekt der Fahrkompetenz erwiesen, der als relevant für das sichere Autofahren gelten kann (z.B. McKenna & Crick, 1994b). Wie man diese sicherheitskritische Fähigkeit angemessen überprüfen kann, muss empirisch untersucht werden. Um Vorhersagen darüber zu treffen, worin eine optimale Operationalisierung der Gefahrenwahrnehmung besteht, können aufgrund des Fehlens spezieller, empirisch gesicherter Modelle zur Fahrexpertise allgemeine Modelle zu Expertise und Expertiseerwerb herangezogen werden. Insbesondere das Konzept der Domänenspezifität erscheint hier zielführend. Expertise zeigt sich deutlicher, je mehr die zu bewältigende Aufgabe den natürlichen Anforderungen, die in einer Domäne bestehen, entspricht (Glaser & Chi, 1988).

In der vorliegenden Arbeit wurden verschiedene Aufgabenformate entwickelt, deren Eignung zur Erfassung von Fahrexpertise durch Experten-Novizen-Vergleiche im Quer- und Längsschnittdesign überprüft wurde. Diejenigen Aufgaben, bei denen deutliche Leistungsunterschiede zugunsten der Experten festgestellt wurden, können als kriteriumsvalide Messungen von Fahrexpertise betrachtet werden. In zwei Vorstudien (N = 268) wurden mithilfe von 2x2-faktoriellen Versuchsdesigns Aufgabenformate getestet, die unterschiedliche Anforderungen an die Teilnehmer stellten (z.B. Mehrfachwahlaufgaben zum verkehrsspezifischen Faktenwissen, Reaktionszeitaufgaben). Neben dem Faktor Expertise (Fahrschüler vs. erfahrene Fahrer) wurde das Präsentationsformat des vorgegebenen Bildmaterials variiert (statisch vs. dynamisch). Um festzustellen, inwieweit die Formate, die sich in den Vorstudien bewährt hatten, auch die Leistungsentwicklung von Fahrschülern abbilden können, wurde eine Längsschnittuntersuchung mit drei Messzeitpunkten (N = 152) durchgeführt.

Insgesamt hat sich gezeigt, dass Gefahrenwahrnehmungsaufgaben im Reaktionszeitaufgabenformat eine deutliche Trennung von Experten und Novizen ermöglichen, während Aufgaben zum verkehrsrelevanten Faktenwissen im Mehrfachwahlaufgabenformat nicht geeignet waren, um zwischen erfahrenen und unerfahrenen Fahrern zu differenzieren. Im Verlauf ihrer Fahrausbildung erreichten die Novizen bezüglich ihres verkehrsspezifischen Wissens das Leistungsniveau der Experten. In den Aufgaben, die im Reaktionszeitaufgabenformat vorgegeben wurden, konnten sich die Teilnehmer über die Zeit nicht steigern. Das Präsentationsformat hatte auf die Validität der Aufgaben keinen generellen Einfluss. Dynamische Präsentationen von Verkehrsszenarien haben die Novizen bei der Bearbeitung von Aufgaben zum deklarativen Faktenwissen jedoch unterstützt.

Vorhersagen, die auf der Grundlage des Konzeptes der Domänenspezifität aus der allgemeinen Expertiseforschung getroffen wurden, konnten bestätigt werden: Wissen und Fähigkeiten werden am besten auf einem anwendungsbezogenen Niveau überprüft. Für die Fahrerlaubnisprüfung bedeutet dies, dass durch den Einsatz handlungsnaher Aufgabenformate eine Validitätssteigerung ermöglicht wird.

Danksagung

Mein erster Dank gilt meinem Betreuer Herrn Prof. Dr. Roland Brünken dafür, dass er mich einerseits stets mit seiner fachlichen Expertise und seinen kreativen Ideen unterstützt hat und mir andererseits viele Entscheidungsfreiheiten, die so wichtig für kreatives Arbeiten sind, gelassen hat. Sein Vertrauen in meine Fähigkeiten hat mich darin bestärkt, mich allen Anforderungen zu stellen.

Mein herzlicher Dank gilt zudem Frau Prof. Dr. Susanne Buch, die besonders in der Anfangsphase meiner Arbeit stets eine gute Ansprechpartnerin für meine zahlreichen methodischen Fragen war.

Bedanken möchte ich mich auch bei Frau Dr. Antje Biermann, die mir in allen Stadien des Projektes als wertvolle Gesprächspartnerin zur Seite gestanden hat. Herrn Tobias Gall und Herrn Patrick Beyer danke ich für die Programmierungsarbeiten im Rahmen meiner Untersuchungen. Ihre Zuverlässigkeit und ihr Mitdenken haben in beträchtlichem Maße zum Gelingen meiner Arbeit beigetragen.

Auch allen studentischen Hilfskräften, die mich im Projekt unterstützt aber auch bei meinen sonstigen Aufgaben entlastet haben, gebührt ein großer Dank. Ich bedanke mich daher besonders bei Julia Kozlik, Sarah Freidinger, Sarah Jähncke und Jessica Müller.

Danken möchte ich auch den Fahrlehrern und Berufsschullehrern, die mir bei der Organisation meiner Erhebungen geholfen haben und natürlich allen rund vierhundert Teilnehmern, die mir ihre Zeit geliehen haben.

Mein Dank gilt auch meiner Familie und insbesondere meinen Eltern für ihre Unterstützung und grenzenlose Zuversicht allen meinen Vorhaben gegenüber.

Die letzten Zeilen möchte ich dem Dank an meinen Ehemann Marc und meine Tochter Nova widmen, die mir Kraft geben, indem sie mich immer daran erinnern, was im Leben am meisten zählt.

Inhaltsverzeichnis

1	EINLEITUNG.....	1
2	THEORETISCHER HINTERGRUND UND LITERATURÜBERBLICK	5
2.1	Expertise	5
2.1.1	Expertise und Kompetenz	6
2.1.2	Experteneigenschaften	8
2.1.3	Entwicklung von Expertise	11
2.1.3.1	Wirkung von Erfahrung auf Wissensorganisation und -repräsentation.....	12
2.1.3.2	Die <i>deliberate-practice</i> -Theorie.....	17
2.1.3.3	Das Fünf-Stufen-Modell des Fertigkeitserwerbs	19
2.1.3.4	Die cognitive-flexibility-Theorie	21
2.1.4	Methoden der Expertiseforschung	22
2.1.4.1	Experten-Novizen-Vergleiche.....	23
2.1.4.2	Längsschnittstudien.....	24
2.1.4.3	Trainingsstudien.....	25
2.1.5	Besonderheiten der Untersuchung von Expertise im Autofahren	26
2.1.6	Zusammenfassung und Forschungsdesiderate	31
2.2	Das Risiko junger Fahranfänger	32
2.2.1	Risikofaktoren.....	34
2.2.1.1	Jugendlichkeitsrisiko.....	34
2.2.1.2	Anfängerrisiko	36
2.2.1.3	Anteile von Jugendlichkeits- und Anfängerrisiko am Gesamtrisiko.....	39
2.2.2	Entwicklung von Fahrkompetenz	40
2.2.3	Zusammenfassung und Forschungsdesiderate	44
2.3	Gefahrenwahrnehmung.....	46
2.3.1	Modelle zur Einordnung der Gefahrenwahrnehmung in den Risikoverhaltensprozess	47
2.3.2	Aufmerksamkeit und das Absuchen der Umwelt in Verkehrssituationen.....	52
2.3.3	Die Validität von Gefahrenwahrnehmungstests.....	56
2.3.3.1	Gefahrenwahrnehmung und Fahrerfahrung	56
2.3.3.2	Gefahrenwahrnehmung und Unfallbeteiligung	57
2.3.4	Methoden zur Erfassung der Gefahrenwahrnehmung und der Einsatz neuer Medien	59
2.3.4.1	Lernen mit statischem vs. dynamischem Bildmaterial	60
2.3.4.2	Testen mit statischem vs. dynamischem Bildmaterial	65
2.3.4.3	Methoden der Erfassung der Gefahrenwahrnehmung	66
2.3.5	Zusammenfassung und Forschungsdesiderate	68

3	PROBLEMSTELLUNG	71
3.1	Erstellung von Aufgaben zur Erfassung von Teilaspekten der Fahrexpertise	71
3.1.1	Kriterium Expertise	71
3.1.2	Auswahl von Inhaltsbereichen	72
3.1.3	Entwicklung von Expertise in fahr- und verkehrsspezifischen Fertigkeiten	73
3.1.4	Identifikation des angemessenen Präsentationsformates (statisch vs. dynamisch)	74
3.2	Empirische Vorgehensweise	76
4	VORSTUDIEN	78
4.1	Vorstudie 1	78
4.1.1	Ziele und Fragestellungen	78
4.1.2	Methode	81
4.1.2.1	Experimentelles Design und Stichprobe	81
4.1.2.2	Versuchsmaterial	84
4.1.2.3	Durchführung	94
4.1.3	Ergebnisse	96
4.1.3.1	Item- und Skalenstatistiken der abhängigen Variablen	96
4.1.3.2	Visualisierungsfähigkeit	99
4.1.3.3	Antwortsicherheit	99
4.1.3.4	Einfluss von Expertise und Präsentationsform auf die Leistung	100
4.1.3.5	Einfluss von Unfallvorgeschichte und kritischen Situationen im Verkehr	104
4.1.4	Diskussion	105
4.1.4.1	Bedeutung der Ergebnisse für die Hypothesen	105
4.1.4.2	Implikationen für das weitere empirische Vorgehen	109
4.2	Vorstudie 2	111
4.2.1	Ziele und Fragestellungen	111
4.2.2	Methode	116
4.2.2.1	Experimentelles Design und Stichprobe	116
4.2.2.2	Versuchsmaterial	118
4.2.2.3	Durchführung	125
4.2.3	Ergebnisse	126
4.2.3.1	Item- und Skalenstatistiken der abhängigen Variablen	126
4.2.3.2	Einfluss von Expertise und Präsentationsform auf die Leistung in den Reaktionsaufgaben	129
4.2.4	Diskussion	137
4.2.4.1	Bedeutung der Ergebnisse für die Hypothesen	137
4.2.4.2	Implikationen für das weitere empirische Vorgehen	141

4.3	Zusammenfassende Diskussion der Vorstudien	143
5	LÄNGSSCHNITTSTUDIE	147
5.1	Ziele und Fragestellungen	147
5.1.1	Ziele	147
5.1.2	Hypothesen	148
5.1.2.1	Hypothesen zu den Experten-Novizen-Vergleichen	148
5.1.2.2	Hypothesen zur Leistungsentwicklung über die Zeit	151
5.2	Methode	153
5.2.1	Experimentelles Design und Stichprobe	153
5.2.2	Versuchsmaterial.....	157
5.2.2.1	Verkehrsspezifische Leistungsmessung (Paralleltests)	157
5.2.2.2	Paper Folding Test	158
5.2.2.3	Hidden Patterns Test	158
5.2.2.4	Persönliche Angaben.....	159
5.2.2.5	Sensation Seeking	160
5.2.3	Durchführung	160
5.3	Ergebnisse	161
5.3.1	Fahrerfahrung und Fahrausbildung	161
5.3.2	Unfälle und kritische Situationen	165
5.3.3	Parallelität der Testformen	166
5.3.4	Kovariaten	167
5.3.5	Ergebnisse der verkehrsspezifischen Leistungsmessung: Experten-Novizen-Vergleiche	169
5.3.5.1	Messzeitpunkt 1	170
5.3.5.2	Messzeitpunkt 2	174
5.3.5.3	Messzeitpunkt 3	177
5.3.6	Ergebnisse der verkehrsspezifischen Leistungsmessung: Leistungsentwicklung über die Zeit.....	181
5.3.6.1	Kovariaten.....	182
5.3.6.2	Verkehrsspezifische Leistungsmaße	182
5.4	Diskussion der Längsschnittstudie	194
5.4.1	Bestätigung der Hypothesen in der Längsschnittstudie.....	194
5.4.1.1	Messzeitpunkt 1	194
5.4.1.2	Messzeitpunkte 2 und 3.....	195
5.4.1.3	Längsschnittstudie.....	196
5.4.2	Zusammenhang von Leistung und räumlichen Fähigkeiten.....	199
6	DISKUSSION UND AUSBLICK.....	201

6.1	Expertise und Expertiseerwerb in der Domäne Autofahren.....	201
6.1.1	Merkmale von Fahrexpertise.....	201
6.1.2	Merkmale des Erwerbs von Fahrexpertise	207
6.2	Bedeutung der Ergebnisse für die Praxis der Fahrerlaubnisprüfung und Fahrausbildung ...	210
6.2.1	Bedeutung für die Gestaltung der Fahrerlaubnisprüfung	210
6.2.2	Bedeutung für die Gestaltung der Fahrausbildung.....	212
LITERATURVERZEICHNIS		215
TABELLENVERZEICHNIS		229
ABBILDUNGSVERZEICHNIS.....		232
ANHANG.....		I

1 Einleitung

Junge Fahranfänger¹ weisen, verglichen mit anderen Personengruppen, zu Beginn ihrer Fahrkarriere ein deutlich erhöhtes Risiko auf, als Fahrer eines PKW im Straßenverkehr zu verunglücken (vgl. z.B. Maycock et al., 1991; Schade, 2001). Im Jahr 2009 gehörten 19 % aller im Straßenverkehr Getöteten sowie jeder fünfte Verletzte der Altersgruppe der 18- bis 24-Jährigen an (Statistisches Bundesamt, 2010a). Ihr Anteil an der Gesamtbevölkerung in Deutschland betrug allerdings lediglich 8,3 %.

Eine Reduktion dieses erhöhten Risikos wird durch unterschiedliche Maßnahmen zu erreichen versucht. Neben der Optimierung der Fahrausbildung - beispielsweise durch die Einführung des Begleiteten Fahrens ab 17 Jahren (BF 17; Stiensmeier-Pelster & Schöne, 2005) - wird auch die Optimierung der theoretischen und praktischen Fahrerlaubnisprüfung beabsichtigt. Bei Maßnahmen zur Fortentwicklung der Fahrerlaubnisprüfung, denen als wichtigstes Ziel die Reduktion des initialen Gefährdungsgipfels direkt nach dem Fahrerlaubniserwerb zugrunde liegt, wird eine bessere Identifikation derjenigen Fahranfänger verfolgt, die aufgrund mangelnder Fahrkompetenz ein besonders hohes Unfallrisiko haben. Um diese Risikopersonen durch die Prüfung zuverlässig von den kompetenten und somit weniger gefährdeten Personen zu unterscheiden, bedarf es geeigneter Prüfungsaufgaben. Bei der Identifikation solcher Prüfungsaufgaben, die eine Trennung zwischen fahrkompetenten und nicht fahrkompetenten Personen erlaubt, handelt es sich um ein Validitätsproblem. Es muss demnach sichergestellt werden, dass die Fahrerlaubnisprüfung Aufgaben enthält, die solche Aspekte der Fahrkompetenz messen, die in einem direkten Zusammenhang mit dem Unfallrisiko stehen. Die deutsche Fahrerlaubnisprüfung für PKW besteht aus einem theoretischen und einem praktischen Teil. In der vorliegenden Arbeit, deren Erhebungen im Zuge der Durchführung des Projektes FE 82.326 (finanziert durch die Bundesanstalt für Straßenwesen, BASt) stattfanden, wird die Verbesserung der theoretischen Fahrerlaubnisprüfung durch valide Aufgaben fokussiert.

In Deutschland wurden unlängst erste Maßnahmen ergriffen, die eine Verbesserung der Güte der theoretischen Fahrerlaubnisprüfung bewirken sollen. So erfolgte im Zeitraum von 2008 bis 2010 eine schrittweise Umstellung der theoretischen Fahrerlaubnisprüfung vom ursprünglichen Papier-Bleistift-Verfahren auf ein computerbasiertes Testsystem. Bei dem Wechsel auf Computertestung wurde eine Adaptation der ursprünglichen Aufgaben

¹ Aus Gründen der Lesbarkeit gilt bei allen personenbezogenen Bezeichnungen die gewählte Form für beide Geschlechter.

vorgenommen, die überwiegend im Mehrfachwahlaufgabenformat vorlagen. Diese Umstellung hat sich vermutlich begünstigend auf die Ökonomie und Standardisierung dieser Prüfung ausgewirkt. Durch eine computerbasierte Testung werden neben Vorteilen, die sich auf die ökonomische und standardisierte Testvorgabe beziehen, auch Aspekte gefördert, die die Validität der Erfassung von Fahrkompetenz betreffen. Die Ablösung einer Papier-Bleistift-basierten Testform bietet die Möglichkeit der Integration innovativer Präsentations- und Aufgabenformate, durch deren Einsatz die Validität der Testung positiv beeinflusst werden könnte. Besonders in der Vorgabe von Testmaterial mit dynamischen Darstellungen von Verkehrsszenarien wird das Potenzial des computerbasierten Testens zur Validitätssteigerung der Prüfung gesehen, da eine Passung zwischen Realität und Testaufgabe hergestellt wird.

Innovative Aufgabenformate, deren Validität empirisch bestätigt wurde, liegen bisher innerhalb der theoretischen Fahrerlaubnisprüfung noch nicht vor. Um solche validen Aufgaben zu erstellen, ist es einerseits notwendig, relevante Inhalte zu identifizieren und andererseits die ausgewählten Inhalte adäquat abzuprüfen. Dieses Problem bezeichnen Horswill und McKenna (2004) als maßgeblich für die Verkehrssicherheitsforschung, deren Bestreben die Bestimmung quantifizierbarer Komponenten der Fahrexpertise ist, die zuverlässig erklären, weshalb einige Fahrer häufiger in Unfälle verwickelt sind als andere. Als Grundlage zur Suche nach Prüfungsinhalten, die an sicherheitsrelevanten Fähigkeiten orientiert sind, dient seit langem die Annahme, dass sich das Risiko der Fahranfänger aus zwei getrennten Komponenten zusammensetzt. Es handelt sich dabei um eine anfängerspezifische Risikokomponente, die sich bei allen Fahranfängern unabhängig von ihrem Alter aufgrund mangelnder Fahrerfahrung zeigt und eine jugendlichkeitsbezogene Risikokomponente, die das Unfallrisiko junger Fahrer beeinflusst (vgl. z.B. Leutner, Brünken & Willmes-Lenz, 2009). Für jede der beiden Risikokomponenten werden spezifische Defizite angenommen, die prinzipiell überprüft werden könnten. Es hat sich gezeigt, dass anfängerspezifische Defizite als schwerwiegender gelten können (Maycock et al., 1991), weshalb man sich in der Prüfung auf die entsprechenden Fähigkeitsbereiche konzentrieren sollte. Defizite von Fahranfängern sind insbesondere in den Bereichen des Absuchens der Verkehrsumwelt auf potenzielle Gefahrenreize und der Aufrechterhaltung der Aufmerksamkeit deutlich nachweisbar (Klauer, S. G., Dingus, Neale, Sudweeks & Ramsey, 2006; McKnight & McKnight, 2003). Es handelt sich dabei um Aspekte der *Gefahrenwahrnehmung*, die in der Forschung konsistent als eine kritische Fähigkeit beim

Autofahren ermittelt wurde (z.B. Biermann, 2007; McKenna & Crick, 1994b; Sümer, Berfu & Birdal, 2007). Im Anschluss an die Identifikation relevanter Inhalte sollte ebenfalls auf der Grundlage empirischer Untersuchungen entschieden werden, wie eine adäquate Messung der Fähigkeiten erfolgen kann. Entscheidend ist dabei auch die Frage nach dem Niveau, auf dem die Lernziele der Ausbildung abgeprüft werden sollen, ob es sich z.B. um deklaratives Faktenwissen handelt oder ob die Lerninhalte flexibel auf unterschiedliche Kontexte angewandt werden sollen. Problematisch sowohl für die Auswahl sicherheitsrelevanter Fähigkeiten als auch für die Bestimmung eines angemessenen Prüfverfahrens ist allerdings die Tatsache, dass empirisch geprüfte Modelle zur Fahrkompetenz und dem Aufbau von Fahrexpertise derzeit noch nicht vorliegen. Aufgrund dessen ist eine geeignete methodische Strategie notwendig, die es erlaubt, ohne ein spezifisches gesichertes Modell Vorhersagen über fahrrelevante Fähigkeiten, deren Entwicklung und Überprüfung zu treffen und zu überprüfen.

Die Frage nach sicherheitsrelevanten Fähigkeiten kann über verschiedene methodische Ansätze beantwortet werden. So stellt z.B. die Analyse von Unfallursachen eine Möglichkeit dar, auf bedeutsame Fähigkeiten zu schließen (Klauer, S. G. et al., 2006). Dabei werden registrierte Unfälle hinsichtlich ihrer Ursachen untersucht und daraus Rückschlüsse auf Defizite in entsprechenden Fähigkeiten der beteiligten Fahrer gezogen. Eine weitere Möglichkeit relevante Fähigkeiten zu identifizieren, besteht im experimentellen Ansatz. Dabei werden Gruppen von Fahrern, die als unterschiedlich in ihrer Fahrkompetenz gelten, hinsichtlich der zu untersuchenden Fähigkeiten verglichen. Es wird dabei angenommen, dass Leistungsunterschiede zwischen den Fahrergruppen auf eine sicherheitsrelevante Fähigkeit hinweisen. Dabei kann sowohl auf Modelle und Befunde als auch auf die Methoden der Expertiseforschung zurückgegriffen werden. Die Forschung zu Expertise und Expertiseerwerb beschäftigt sich klassischerweise mit der überdurchschnittlichen Leistung einzelner Personen (Experten) in einer bestimmten Domäne (Gruber, 1994). Befunde dieser Forschungsrichtung lassen sich aber auch auf Prozesse beim Aufbau alltäglicher Fähigkeiten wie z.B. dem Autofahren anwenden. Außerdem werden Methoden der Expertiseforschung bei der Validierung von Verfahren zur Leistungsmessung eingesetzt.

In der vorliegenden Arbeit werden quasi-experimentelle Studien vorgestellt, die sich mit Expertise, Expertiseerwerb und der Erfassung von Expertise in der Domäne Autofahren befassen. Der theoretische Rahmen für die Erstellung und Überprüfung der neuen Aufgabenformate umfasst drei Forschungsbereiche, die in einer Reihenfolge mit

absteigendem Abstraktionsgrad dargestellt werden (Kapitel 2). Zunächst wird auf Merkmale von Expertise und Expertiseentwicklung allgemein eingegangen. Bisher gibt es kein konkretes empirisch gesichertes Modell zur Erklärung des Erwerbs von Fahrexpertise. Konzepte aus der Expertiseforschung werden daher auf die Domäne Autofahren bezogen, wobei Schlüsse auf ihre Besonderheiten und angemessene Methoden zur Bestimmung der Fahrexpertise gezogen werden. Im Anschluss daran werden Aspekte von Fahrexpertise im Einzelnen anhand von Risikofaktoren von Fahrnovizen aufgeschlüsselt. Dabei wird darauf eingegangen, in welchen Merkmalen erfahrene Fahrer sich von unerfahrenen Fahrern unterscheiden und inwieweit diese Merkmale sicherheitsrelevant sind. Eines dieser Merkmale hat sich in der Forschung besonders konsistent als Kriterium für Fahrexpertise erwiesen: die *Gefahrenwahrnehmung*, die den dritten Forschungsbereich darstellt. Es handelt sich dabei um die Fähigkeit, potenziell gefährliche Umweltreize schnell zu identifizieren (z.B. McKenna & Crick, 1991). Anschließend werden die Ergebnisse der Literaturrecherche zusammengefasst und Forschungsdesiderate formuliert (Kapitel 3). Als angemessene Methode zur Untersuchung der Eignung verschiedener verkehrsspezifischer Aufgabenformate zur Erfassung von Fahrexpertise wird eine Kombination von querschnittlich und längsschnittlich geplanten Experten-Novizen-Vergleichen erachtet. Im folgenden Kapitel (Kapitel 4) werden zunächst zwei Vorstudien vorgestellt, die mithilfe von Experten-Novizen-Vergleichen die Kriteriumsvalidität verschiedener Aufgabentypen zur Messung von Verkehrswissen und Gefahrenwahrnehmung bestimmen. Neben der Expertise (Experten vs. Novizen) wird der Faktor Präsentationsformat des Aufgabenmaterials (statisch vs. dynamisch) experimentell variiert. Geeignete Aufgabenformate gehen in die anschließende Längsschnittuntersuchung mit drei Messzeitpunkten ein. Dabei wird die Leistungsentwicklung in den eingesetzten Aufgaben von Fahrschülern und erfahrenen Fahrern gegenüber gestellt (Kapitel 5). Die Ergebnisse werden im Anschluss hinsichtlich ihrer Bedeutung für das Verständnis der Beschaffenheit und Entwicklung von Fahrexpertise sowie für einen Einsatz von dynamischen Präsentationsformaten und innovativen Antwortformaten in der theoretischen Fahrerlaubnisprüfung diskutiert (Kapitel 5).

2 Theoretischer Hintergrund und Literaturüberblick

2.1 Expertise

Spätestens seit den Studien von Adriaan D. de Groot (z.B. de Groot, 1965) sowie William G. Chase und Herbert A. Simon (z.B. Chase & Simon, 1973) Ende der 60er Jahre und Anfang der 70er Jahre wird der Expertiseforschung erhebliche Bedeutung in der kognitiven Psychologie beigemessen. Die Expertiseforschung beschäftigt sich in der Regel mit außerordentlichen Leistungen, die einzelne Personen in einem speziellen Inhaltsbereich erbringen und mit der Frage, wie ein solch hohes Leistungsniveau erreicht werden kann (Gruber, 1994). Expertise kann zudem als maximale Anpassung an Anforderungen der Umwelt verstanden werden, wobei die Erforschung der Entwicklung von Expertise auch wichtige Einsichten in generelle menschliche Anpassungsmechanismen ermöglichen kann (Ericsson & Lehmann, 1996; Ericsson & Ward, 2007). Besonders die Lern- und Instruktionspsychologie profitiert von den Erkenntnissen über den Einfluss von Vorwissen – als einen Aspekt von Expertise – auf die Leistung einer Person und deren Wechselwirkungen mit Gestaltungsmerkmalen von Lernmaterial (z.B. Sweller & Chandler, 1994). Mittlerweile spielt somit das Konzept der Expertise eine bedeutsame Rolle in Bezug auf die Analyse von Leistung und Verhaltensbeständigkeit, in der Aufgabenanalyse, in Lern- und Trainingsstudien, bei der Gestaltung von Benutzeroberflächen sowie bei der Modellierung kognitiver Prozesse in den unterschiedlichsten Inhaltsbereichen (Farrington-Darby & Wilson, 2006). In den Anfängen der Expertiseforschung stellten jedoch Spiele und insbesondere das Schachspiel die zentrale Domäne dar, in der herausragende Leistungen untersucht wurden (Chase & Simon, 1973; de Groot, 1965). Schach wird daher zuweilen – in Anlehnung an die zahlreichen genetisch-biologischen Arbeiten zur Taufleie – als die *Drosophila* der Expertiseforschung bezeichnet (z.B. Gruber, 1994; Gruber & Lehman, 2007; Simon & Chase, 1973). In einer Überblicksarbeit stellten Bédard und Chi (1992) Ergebnisse bisheriger Studien zur Expertise dar. Der Abriss der Autoren verdeutlicht, dass bis zu diesem Zeitpunkt hauptsächlich Experten-Novizen-Vergleiche im Labor durchgeführt wurden, die in der Regel aus Gedächtnisexperimenten und Aufgaben zur Problemkategorisierung aus Einführungskursen an der Universität z.B. im Fach Physik bestanden. Als globale Grundlage für Expertise in einem bestimmten Bereich – in einer sogenannten *Domäne* – wurde eine enorme, gut organisierte domänenspezifische Wissensbasis identifiziert. Die Autoren bezweifelten dabei die Generalisierbarkeit der Ergebnisse auf Expertise im realen Leben.

Um diese Lücke zu füllen, wurden später Fähigkeiten von Experten im Zusammenhang mit dem Umgang mit realistischen Problemen, die den Aufgaben in einer spezifischen Domäne eher entsprachen, untersucht. Dabei wurden Personen einbezogen, die Expertise in unterschiedlichen Bereichen erworben hatten. Im Zentrum der Expertiseforschung standen hierbei insbesondere semantisch reichere, also wissensabhängige Domänen (Krems, 1994), wie z. B. medizinische Diagnostik (z.B. Lesgold et al., 1988; Rikers, Schmidt & Moulaert, 2005; Schmidt, H. G. & Boshuizen, 1993), Lehre und Unterricht (z.B. Bromme, 1992) oder Physik (z.B. Chi, Feltovich & Glaser, 1981; Friege, 2000).

2.1.1 Expertise und Kompetenz

Die Begriffe Expertise und Experte werden sowohl im Alltag als auch in der psychologischen Forschung gebraucht, wobei die Bedeutung der Begriffe in beiden Bereichen in weiten Teilen übereinstimmt. Im Alltagsgebrauch bezeichnet man einen Menschen als Experten, der sich auf einem bestimmten Gebiet sehr gut auskennt, auf diesem Gebiet Probleme effektiv lösen kann und dem man dabei vertrauen kann (Gruber, 1994). Krems (1994) stellt folgende drei Merkmale heraus, die Experten aus alltagspsychologischer Sicht auszeichnen: Effizienz, bereichsspezifisches Wissen und Können sowie Erfahrung. Experten werden demnach als effektiver als Novizen betrachtet. Das bedeutet, dass sie innerhalb eines bestimmten Zeitraums überdurchschnittlich viele Aufgaben mit besonders wenig Aufwand bewältigen. Außerdem können Experten bei der Bewältigung von spezifischen Aufgaben einerseits auf ein umfangreiches und differenziertes Fachwissen in einem speziellen Sachgebiet und andererseits auf Methoden und Strategien zur Bewältigung zurückgreifen. Um dieses Wissen und Können auf Expertenniveau zu erreichen, bedarf es langjähriger, intensiver Beschäftigung mit einem bestimmten Sachgebiet.

Im psychologisch-wissenschaftlichen Sinne lieferte Posner (1988) in seinem Buchkapitel „*What it is to be an expert*“ eine Definition des Expertenbegriffs, die zum einen als eher unspezifisch, zum anderen aber als die am weitesten verbreitete Definition bezeichnet werden kann. Eine Zusammenfassung dieser Definition wurde von Gruber (1994) im Deutschen folgendermaßen formuliert: „Ein Experte ist eine Person, die auf einem bestimmten Gebiet dauerhaft (also nicht zufällig oder singulär) herausragende Leistung erbringt“ (S. 10).

Expertise kann sich demnach in einer herausragenden Performanz in einem von höchst unterschiedlichen Bereichen wie z.B. bei manuellen Fertigkeiten, im mentalen oder auch im künstlerischen Bereich zeigen (Gruber, 1994).

Zahlreiche Konzepte der Lehr-, Lern- und Bildungsforschung sind dem der Expertise ähnlich. An dieser Stelle wird lediglich auf den bedeutungsverwandten Begriff *Kompetenz* näher eingegangen. Die beiden Begriffe Kompetenz und Expertise werden in der Alltagssprache sowie in wissenschaftlichen Artikeln häufig als Synonyme gebraucht. Explizit unterschieden werden die Begriffe selten. Das Konzept der Kompetenz wird hier näher beschrieben, da es neben der Tatsache, dass es sich in der deutschen Alltagssprache im Sinne von „etwas können“ um einen relativ geläufigen Begriff handelt, derzeit in der Bildungsforschung allgegenwärtig erscheint. Die Datenbank PsycInfo zeigt für die Begriffe *competence*, *competencies* und *competency*, dass die Verwendung des Begriffes – und damit wahrscheinlich auch die Bedeutung des Konzeptes für die Forschung – in den letzten zehn Jahren deutlich angestiegen zu sein scheint. Waren es zwischen 1985 und dem Jahre 1999 durchschnittlich etwa drei Veröffentlichungen pro Tag zum Thema Kompetenz, ist die Zahl der täglichen Veröffentlichungen seit der ersten PISA Studie im Jahre 2000 auf sechs bis sieben Veröffentlichungen gestiegen.

Der Begriff Kompetenz wird zumeist eingesetzt, wenn Leistungsfähigkeit aus einer anwendungsbezogenen Perspektive betrachtet wird. So wird er in der Regel zur Beschreibung verwendet, inwieweit eine Person die Anforderungen, die in einem bestimmten Kontext gestellt werden, erfüllen kann (Klieme, Maag-Merki & Hartig, 2007). In der aktuellen Diskussion hat sich eine Kompetenzdefinition von Weinert (2001) etabliert, wonach Kompetenz verstanden wird als

die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösung in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können. (S. 27, f).

Kompetenzentwicklung umfasst demnach einerseits den Aufbau von Wissen und Können. Die Definition schließt aber andererseits auch ein, dass diese erworbene kognitive oder motorische Basis in angemessener Weise in eine Handlung umgesetzt werden kann und dass auch die Bereitschaft dazu besteht. Der Begriff Kompetenz wird häufig dazu verwendet, um bestimmte Bildungsziele zu formulieren (Hartig & Klieme, 2006). Die Autoren charakterisieren Kompetenzen allgemein als bereichsspezifisch, wobei die Fähigkeit zu Transfer und Verallgemeinerungen eingeschlossen wird. Es handelt sich bei Kompetenzen um Kenntnisse, Fertigkeiten oder Routinen, die durch Lernen erworben werden können oder sogar müssen (Hartig & Klieme, 2006).

Eine Verbindung zwischen pädagogischen Basiskonstrukten, psychologischen Kompetenzmodellen und entsprechenden Messverfahren stellt sich zuweilen schwierig dar, ist aber für die Kompetenzdiagnostik unverzichtbar (z.B. Hartig & Klieme, 2006). Sollen spezifische Kompetenzen empirisch untersucht werden, wie es z.B. in den Vergleichsstudien der OECD bei repräsentativen Schulleistungsstudien wie PISA (Baumert et al., 2001) angestrebt wird, muss man individuelle Kompetenzausprägungen möglichst eindeutig bestimmen. Dabei können Kompetenzen in Teilkompetenzen zerlegt werden. Diese Strukturierung erfolgt laut Hartig und Klieme (2006) anhand der unterschiedlichen Anforderungen, die zur Lösung verschiedener Aufgaben bewältigt werden müssen. Zur Diagnostik von Kompetenzen werden in der Regel standardisierte Tests, ähnlich denen der psychologischen Leistungsdiagnostik, herangezogen (Hartig & Klieme, 2006). Die Autoren stellen zudem dar, dass die Konstruktion dieser Instrumente auf theoretischen Modellen basiert, die sich an zwei Formen von Kompetenzmodellen orientieren können: Strukturmodelle, die Kompetenzdimensionen aufspannen oder Niveaumodelle, welche in der Lage sind, spezifische Fähigkeiten zu beschreiben, die unterschiedliche Kompetenzniveaus charakterisieren. Kompetenzmodelle werden jeweils für unterschiedliche Bereiche (meist Schulfächer) getrennt entwickelt (Klieme et al., 2007).

Die beiden Konzepte Expertise und Kompetenz stimmen in ihren Hauptprinzipien überein: Sie werden als bereichsspezifisch und erlernbar angenommen. Die Begriffe Expertise und Kompetenz unterscheiden sich daher weniger auf inhaltlicher, sondern eher auf konzeptueller Ebene. Während die Charakterisierung und Diagnostik spezifischer Kompetenzen auf explizit für sie entwickelten domänen- oder fachspezifischen hierarchischen Kompetenzmodellen basiert, geben die meisten Expertise- und Expertiseerwerbsmodelle vor, generalisierbar zu sein. Dies bietet die Möglichkeit, vor dem theoretischen Hintergrund von Modellen zur Expertise und zum Expertiseerwerb, immer neue Domänen zu untersuchen, für die bisher noch keine Kompetenzmodelle vorliegen. Das Ziel dieser Arbeit ist die Untersuchung von Fähigkeiten in der Domäne Autofahren, für die bisher noch kein Kompetenzmodell vorliegt. Daher basiert die Forschungsmethodik zur Untersuchung der vorliegenden Fragestellungen auf theoretischen Modellen und Forschungsergebnissen aus dem Bereich der Expertiseforschung.

2.1.2 Experteneigenschaften

Glaser, Chi und Farr (1988) haben die Ergebnisse der Anfänge der Expertiseforschung zusammengeführt und geben einen Überblick darüber, durch welche Eigenschaften sich

Experten auszeichnen. Die Autoren stellen zusammenfassend sieben Schlüsselcharakteristiken von Experten dar, die sich in der Forschung als robust und über verschiedene Domänen generalisierbar erwiesen haben.

1. Eine wichtige Erkenntnis ist, dass Experten hauptsächlich auf ihrem Spezialgebiet herausragende Leistungen erbringen. Ihre Leistungen stechen demnach besonders hervor, wenn es um den Bereich geht, in dem sie viel Wissen und Fähigkeiten erworben haben. Dieser Vorteil besteht nicht unbedingt in anderen Domänen. Es findet also kein Transfer von Expertise statt. Diese Erkenntnis wurde z.B. durch die Forschungsergebnisse von Chase und Simon (1973) gewonnen. Dabei zeigte sich, dass Schachexperten beim Erinnern von bedeutungsvollen Schachkonstellationen den Schachnovizen deutlich überlegen waren. Dieser Unterschied zeigte sich jedoch nicht für das Erinnern zufälliger Schachkonstellationen.
2. Außerdem nehmen Experten auf ihrem Gebiet große bedeutungsvolle Muster beziehungsweise Zusammenhänge wahr. So wird angenommen, dass sich Schachmeister bedeutungsvolle Schachkonstellationen besser merken können, weil sie diese als Ganzes wahrnehmen. Dieses Ergebnis konnte auch für andere Domänen, wie z.B. auf dem Gebiet der Interpretation von Röntgenbildern, gezeigt werden (Lesgold et al., 1988).
3. Ein weiterer Forschungsbefund ist, dass Experten auf ihrem Gebiet schneller spezifische Probleme lösen und weniger Fehler machen als Novizen. Dies wird im Falle einfacher Tätigkeiten, wie Schreibmaschineschreiben, darauf zurückgeführt, dass durch viel Übung Handlungen automatisiert werden und damit Verarbeitungskapazität für andere Aspekte des Problemlösens frei wird (Gentner, 1988).
4. Es hat sich zudem gezeigt, dass Experten mehr Informationen aus ihrem Spezialgebiet in Kurzzeit- und Langzeitgedächtnis speichern können. Chase und Ericsson (1982) konnten beides für das Erinnern von Zahlenreihen eines trainierten Experten bestätigen.
5. Außerdem nutzen Experten grundlegende Prinzipien zur Verarbeitung domänenspezifischer Probleme. Novizen hingegen bearbeiten Probleme eher auf einem oberflächlichen Niveau. Chi, Feltovich und Glaser (1981) konnten diese Annahme

bestätigen, indem sie in verschiedenen experimentellen Studien die Art der Kategorisierung von Problemen in der Domäne Physik durch Experten und Novizen verglichen. Es zeigte sich dabei, dass die gebildeten Kategorien der Experten sich qualitativ von denen der Novizen unterschieden. Die auf Prinzipien (z.B. Einbezug aller physikalischen Gesetze) beruhende Kategorisierung der Experten erlaubte es ihnen, die Anzahl der möglichen Lösungswege auf ein Minimum zu beschränken und damit effektiv weiterzuarbeiten.

6. Eine weitere Eigenschaft von Experten stellt die Tendenz dar, sich beim Lösen eines Problems relativ zur Gesamtzeit, die sie für die Lösung eines Problems benötigen, lange mit der qualitativen Analyse des Problems zu beschäftigen. Novizen hingegen analysieren das Problem mit seinen spezifischen Merkmalen eher oberflächlich und widmen diesem Aspekt des Problemlösens verhältnismäßig wenig Zeit. Experten versuchen in der Anfangsphase zunächst das Problem zu verstehen, während Novizen schneller dazu übergehen, verschiedene Lösungswege auszuprobieren (Glaser & Chi, 1988). Chi, Feltovich und Glaser (1981) erklären dies damit, dass Experten beim Problemlösen eine *Top-Down-Strategie* einsetzen, während umgekehrt die Novizen zur Problemlösung eine *Bottom-Up-Strategie* nutzen. Bei Experten wird demnach zur Problemlösung zunächst ihr deklaratives und prozedurales Wissen aktiviert, während dies bei Novizen nicht erfolgt.

7. Zuletzt ließ sich zeigen, dass Experten über gute Fähigkeiten zum *Selbst-Monitoring* verfügen und daher die eigenen Fähigkeiten gut einschätzen können. So konnten Chi, Glaser und Rees (1982) zeigen, dass Experten zuverlässiger als Novizen vorhersagen konnten, welche Physikprobleme schwieriger zu lösen sein würden als andere. Neben der Einschätzung der globalen Schwierigkeit der jeweiligen Aufgaben sollten die Teilnehmer außerdem markieren, welche Begriffe bzw. Satzteile in der Aufgabenstellung sie als relevant zur Einschätzung der Schwierigkeit betrachteten. Dabei zeigte sich, dass Experten und Novizen zu sehr ähnlichen Ergebnissen kamen. Die Autoren vertreten daher die Auffassung, dass Novizen zwar die wichtigen Hinweise in der Aufgabenstellung erkennen, sie jedoch im Gegensatz zu den Experten Probleme damit haben, Folgerungen zu ziehen, die sich für Aspekte ergeben, die nicht explizit in der Aufgabenstellung genannt werden.

Gruber und Mandl (1995) führen als Kriterien von Expertise an, dass Experten über eine große Wissensbasis verfügen, reichlich domänenspezifische praktische Erfahrung haben und erfolgreich spezifische Probleme erkennen und lösen können. Ihre Handlungen kontrollieren Experten auf einer höheren (metakognitiven) Ebene und sie agieren effizient, ohne Fehler und sehr genau. Experten sind zudem in der Lage, auf neue Probleme flexibel zu reagieren. Diese Kriterien unterscheiden sich nicht deutlich von den Experteneigenschaften, die in der früheren Literaturstudie von Glaser, Chi und Farr (1988) herausgestellt wurden. Lediglich der flexible Umgang der Experten mit neuen Problemen wurde in früheren Arbeiten nicht so deutlich herausgestellt. Vermutlich hat dieser Aspekt an Bedeutung gewonnen, nachdem Expertiseforschung zunehmend auch in komplexeren Domänen durchgeführt wurde, in denen vielfältigere Anforderungen gestellt werden und demnach der Umgang mit diesen beobachtet werden konnte. In einer späteren Überblicksarbeit von Cellier, Eyrolle und Marine (1997) wurde als ein weiteres Expertisemerkmal herausgestellt, dass Experten bessere Voraussagen treffen können. Sie antizipieren Schwierigkeiten und reagieren präventiv statt reaktiv auf Störungen und Probleme.

Überblicksarbeiten, wie z.B. die von Glaser, Chi und Farr (1988) fassen die Befunde über unterschiedliche Domänen zusammen. Es wird dabei kein Hinweis darauf gegeben, dass sich die grundlegenden Eigenschaften von Experten in unterschiedlichen Domänen deutlich voneinander unterscheiden.

2.1.3 Entwicklung von Expertise

Während die Ansichten über die Eigenschaften eines Experten in der Fachwelt weitestgehend übereinstimmen, herrscht Uneinigkeit darüber, wie Expertise ausgebildet wird (Rikers & Paas, 2005). Eine frühe Ansicht, die typischerweise mit Galtons Arbeit (1969/1979) in Verbindung gebracht wird, besagt, dass grundlegend für den Erwerb von Expertise einzelne angeborene Unterschiede in generellen Fähigkeiten sind, die nicht durch Training oder Erfahrung verändert werden können. Expertise wird in manchen Domänen, wie z.B. Musik oder Sport heute noch als ein eher dispositionales Phänomen aufgefasst (Gruber, 1994; Gruber & Lehman, 2007). Begabung wird dabei als Grundlage dafür gesehen, dass hohe Leistungsfähigkeit überhaupt erbracht werden kann. Die Expertiseforschung im lehrerpsychologischen Bereich stützt sich aber in der Regel auf herausragende domänenspezifische Leistungen mit Inhalten, die durch Erfahrung und Training erlernbar sind. Ericsson und Smith (1991b) betonen sogar, dass das Besondere an der Expertiseforschung ist, dass hypothetische Konstrukte wie z.B. Intelligenz als

Bedingungsfaktoren herausragender Leistungen nicht einbezogen werden. Ericsson, Krampe und Tesch-Römer (1993) fassten zusammen, dass die Suche nach genetischen Merkmalen, die Expertise voraussagen, nicht erfolgreich war. Eine Ausnahme bilde dabei lediglich die Körpergröße bei einigen Sportarten. Chase & Simon (1973) argumentieren, Expertise sei in vielen Domänen das Ergebnis vieler Jahre Erfahrung und des Aufbaus einer umfangreichen Wissensbasis. Dabei fanden die Autoren, dass es mindestens zehnjähriger Übung bedarf, bevor ein Schachspieler internationales Niveau erreicht hat. Diese *10-Jahres-Regel* konnte auch für andere Domänen bestätigt werden (für einen Überblick vgl. Ericsson et al., 1993).

Darüber, wie sich der Expertiseerwerb im Einzelnen vollzieht, bestehen unterschiedliche Vorstellungen. Die Ursache für diese Uneinigkeit könnte in der Diversität der Domänen liegen, in denen Expertiseerwerb stattfindet. Je nach Domäne scheinen für den Expertiseerwerb entweder dispositionale oder erfahrungsabhängige Faktoren stärker verantwortlich zu sein (Gruber, 1994). Empirische Evidenz liegt, mit wenigen Ausnahmen, eher für erfahrungsbezogene Faktoren des Erwerbs von Expertise vor. Die Anlage kann daher nicht als eine für den Expertiseerwerb hinreichende Bedingung erachtet werden. Notwendig scheint hingegen die Erfahrung im betreffenden Bereich für die Ausbildung von Expertise darin zu sein. Im Folgenden sollen daher ausschließlich Modelle vorgestellt werden, denen Erfahrung und Übung in einer bestimmten Domäne als Hauptursachen für den Expertiseerwerb zugrunde liegen.

2.1.3.1 Wirkung von Erfahrung auf Wissensorganisation und -repräsentation

Experten verfügen in ihrem Fachgebiet in der Regel über sehr viel Wissen (Posner, 1988). Diese Erkenntnis veranlasste Expertiseforscher zunächst zu der Annahme, dass Expertiseerwerb als Wissensakkumulation in einem besonderen Umfang nur erfolgen kann, wenn eine Person über eine anlagebedingt außerordentlich hohe Gedächtniskapazität verfügt (vgl. Gobet, 1996). Tatsächlich ist der stabilste Befund aus der Expertiseforschung, dass Experten sich besser an Objekte oder Konstellationen in ihrer Domäne erinnern können, als Personen, die auf diesem Gebiet noch Neulinge sind (Gruber & Mandl, 1995). Bekannt sind besonders Schachexperimente, in denen sich Schachmeister und Schachnovizen merken mussten, welche Schachfiguren in welchem Muster auf einem Schachbrett verteilt waren (Chase & Simon, 1973; Simon & Chase, 1973). Beim Spielen selbst ist dies allerdings keine Aufgabe, die zu einem Zeitpunkt im Spielgeschehen relevant wäre.

Während man zunächst Anlagefaktoren als hauptverantwortlich für außergewöhnliche Gedächtnisleistungen hielt, zeigte sich später, dass die erhöhte Speicherkapazität im

Gedächtnis von Experten mit einer effektiveren Wissensorganisation zusammenhängt. In ihrer Überblicksarbeit fassten Bédard und Chi (1992) zusammen, dass sich das Expertenwissen nicht nur quantitativ von dem Wissen der Novizen unterscheidet. Wesentlicher erscheint den Autoren der Aspekt, dass Experten ihr Wissen auf eine Art gespeichert haben, die ihr Wissen nutzbarer, erreichbarer und somit effizienter macht. Erklären kann diese Erkenntnis unter anderem die so genannte *Chunking*-Theorie (Chase & Simon, 1973; Simon & Chase, 1973). Die *Chunking*-Theorie gilt als eine allgemeine Theorie, das heißt, sie ist über unterschiedliche Domänen generalisierbar und gilt zudem nicht nur für Experten, sondern auch andere Personen verfügen über die postulierten Mechanismen. Der Begriff *Chunk* wurde durch Miller (1956) eingeführt, der zeigen konnte, dass die Kapazität des Kurzzeitgedächtnisses auf 7 (+/- 2) Einheiten begrenzt ist, die Informationsmenge einer einzelnen Einheit für sich jedoch nicht beschränkt ist. Miller ging davon aus, dass durch Übung Wissens Elemente nicht mehr einzeln abgespeichert, abgerufen und verarbeitet werden, sondern zu immer größer werdenden Einheiten, den sogenannten *Chunks*, zusammengefasst werden. Diese Zusammenfassung funktioniert nur dann effektiv, wenn im Langzeitgedächtnis Informationen abgespeichert sind, mit deren Hilfe neue Informationen strukturiert werden können (Solso, 2005). Experten haben einen Erinnerungsvorteil, weil sie aufgrund ihres domänenspezifischen Wissens größere *Chunks* bilden können (Krems, 1994). Demnach sind Personen, die ihr Wissen organisiert und strukturiert haben in der Lage größere Informationsmengen zu erinnern als Personen, die über eine solche Wissensstruktur nicht verfügen.

Ergänzt wird die *Chunking*-Theorie durch den *Pattern Recognition*-Ansatz von Chase und Simon (1973), die am Beispiel des Schachspiels belegten, dass Experten über *Chunks* in Form von Schachmustern in ihrem Langzeitgedächtnis verfügten. Ein Beleg dafür ist, dass bedeutungsvolle Konstellationen von Schachexperten besser erinnert wurden als Konstellationen, die keine besondere Bedeutung hatten. Es wurde vermutet, dass die Experten auf bedeutungsvolle Schachaufstellungen über Labels schnell zugreifen können. Dies unterstützten auch bereits Befunde von De Groot (1965), der zeigen konnte, dass Schachexperten im Gegensatz zu Schachspielern mit einem geringeren Expertisegrad sehr schnell mit dem optimalen Zug auf eine neu dargebotene Schachkonstellation reagieren konnten. De Groot schloss daraus, dass die Unterschiede zwischen Schachspielern auf unterschiedlichen Leistungsniveaus auf Prozessen innerhalb der ersten Phase, nachdem eine neue Konstellation der Figuren auf dem Schachbrett erscheint, basieren. Diese schnelle adäquate Reaktion setzt voraus, dass Schachkonstellationen zusammen mit der passenden

Reaktion darauf abgespeichert und für die Experten verfügbar vorliegen (Chase & Simon, 1973).

In den vorgestellten Theorien beruht der Erwerb von Expertise darauf, dass quantitativ mehr Wissen gespeichert werden kann. Es hat sich jedoch gezeigt, dass Experten Informationen auch qualitativ anders verarbeiten. Die *Skilled-Memory-Theorie* (Chase & Ericsson, 1982; Ericsson & Staszewski, 1989) besagt, dass Experten Novizen deshalb übertreffen, weil sie in der Lage sind, ihr Langzeitgedächtnis effizienter zu strukturieren. Neue Informationen werden mit bestehenden semantischen Wissensmustern im Langzeitgedächtnis verknüpft. Die gespeicherte Information kann schnell durch sogenannte *Cues*, die bei der Enkodierung der Information assoziiert wurden, abgerufen werden. Dieser Abruf von Informationen geschieht bei Experten in differenziert ausgearbeiteten Form, da die *Cues* mit hierarchischen Strukturen verbunden sind und es erlauben, mit einem einzigen Schlüsselreiz ein ganzes Netz von Informationen abzurufen. Die enkodierte Information wird im Langzeitgedächtnis gespeichert und kann demnach auch nach einer längeren Zeitspanne noch abgerufen werden.

Erweitert und verallgemeinert wurde die *Skilled Memory-Theory* durch den *Long-Term-Working-Memory-Ansatz* von Ericsson und Kintsch (1995). Dabei wird eine Ausdehnung der traditionellen Konzepte des Arbeitsgedächtnisses, die den Gebrauch temporärer Speicher einschließen, auf ein Konzept des Arbeitsgedächtnisses vorgenommen, das auch auf im Langzeitgedächtnis gespeicherte Informationen in hoch effizienter Form zugreifen kann.

Die bisher vorgestellten Theorien beschäftigen sich mit dem Aufbau von Wissensstrukturen und dem Abruf von Wissen auf der Basis dieser Strukturen. Dabei wird die Art des Wissens nicht genauer spezifiziert. Es kann jedoch angenommen werden, dass sich die Art des Wissens zwischen verschiedenen Domänen unterscheidet und damit eventuell auch der Erwerb und Abruf des Wissens.

Andersons Theorie zum Fertigkeitserwerb, der als theoretischer Rahmen das *Adaptive-Control-of-Thought-Modell* (kurz *ACT**) zugrunde liegt, besagt, dass beim Erwerb von Expertise ein Wandel in der Wissensrepräsentation stattfindet. Anderson unterscheidet drei Stadien des Fertigkeitserwerbs (1982). In der deklarativen Phase verfügt der Lerner lediglich über domänenspezifisches Faktenwissen. In der prozeduralen Phase wird das deklarative Fachwissen direkt in Form von Vorgängen für das Ausüben der Fertigkeit ausgedrückt. Der Übergang von der deklarativen Phase zur prozeduralen Phase erfolgt in einer Phase der

Wissenskompilierung, die durch den wiederholten Einsatz des Wissens erfolgt. Problemlösen erfolgt der Theorie zufolge nach wenn-dann-Regeln, die durch Erfahrung immer besser ausgebildet werden. Expertenwissen sollte dieser Theorie zufolge in prozeduraler Form vorliegen. Anwendbar ist diese Theorie auf das Lernen von Routinehandlungen, bei denen Wissen zum Teil nicht mehr bewusst abrufbar ist (Anderson, 1987). Routinehandlungen liegen bei Experten als automatisierte Prozesse vor und verbrauchen keinerlei kognitive Ressourcen, sie werden nicht von Nebenaufgaben gestört und stören diese auch nicht (Kahneman & Treisman, 1984). Expertise bei Routinehandlungen kann auf der Grundlage des Informationsverarbeitungsmodells von Atkinson und Shiffrin (1969) erklärt werden (Leutner & Brünken, 2002). Diesem Modell zufolge finden zunächst Wahrnehmungsprozesse statt. Die wahrgenommene Information gelangt daraufhin in den Arbeitsspeicher und wird danach im Langzeitgedächtnis gespeichert (vgl. Abbildung 2.1). Im Langzeitspeicher abgelegte Informationen können im Arbeitsspeicher mit aktuellen Informationen abgeglichen werden. Aufmerksamkeitsprozesse, emotionale und motivationale Prozesse sowie metakognitive Kontrollprozesse wirken sowohl auf die Wahrnehmung als auch auf die Verarbeitung im Arbeits- und Langzeitspeicher ein. Wenn eine Reaktion notwendig ist, werden psychomotorische Prozesse in Gang gesetzt, die sich in sichtbarem zielgerichtetem Verhalten niederschlagen. Eine Ausnahme bilden Vorgänge, die automatisiert ablaufen. Der Vorgang bei automatisierten Prozessen kann in Abbildung 2.1 nachvollzogen werden, indem man von den Wahrnehmungsprozessen ausgehend dem Pfeil unmittelbar zu den psychomotorischen Prozessen folgt, die entsprechendes Verhalten aktivieren. Diese unmittelbaren Reaktionen sind durch ein hohes Maß an Übung zu Routinehandlungen geworden.

Automatisierte Reaktionen zeichnen sich dadurch aus, dass sie schneller erfolgen und weniger fehleranfällig sind als bewusst gesteuerte Prozesse. Nicht immer zeigt sich prozeduralisiertes Wissen jedoch in automatischen Routinehandlungen. Die Prozeduralisierung von Wissen kann auch zu einer gesteigerten Problemlösefähigkeit führen.

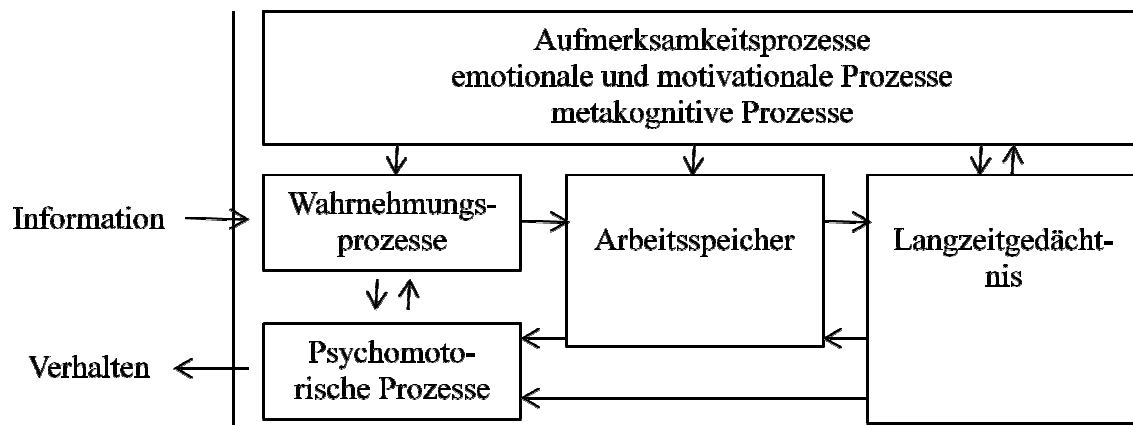


Abbildung 2.1

Modell der Informationsverarbeitung (nach Atkinson & Shiffrin, 1969; Shiffrin & Atkinson, 1971, aus Leutner & Brünken, 2002, S.77)

Die vorangegangenen Theorien können den Expertiserwerb durch Erfahrung und Übung in wenig komplexen Domänen erklären. Andere Konzepte veranschaulichen den Aufbau von Expertise in komplexeren Inhaltsbereichen.

Kolodner (1983) zufolge wird im Schulunterricht sowie in den Vorlesungen an den Hochschulen überwiegend deklaratives Wissen erworben. Die praktische Umsetzung und Anwendung dieses Wissens gestaltet sich für Schüler und Studenten trotz der beträchtlichen Wissensmenge über die sie verfügen noch schwer. Später wenn sie im Beruf tagtäglich aus ihrem Wissen Problemlösungen ableiten müssen, fällt ihnen dieser flexible Umgang mit ihrem deklarativen Wissen zunehmend leichter. Es wird deutlich, dass man beim Übergang zum Experten nicht nur immer mehr Wissen anreichert, sondern auch lernt, das Wissen effektiver anzuwenden. Vermittelt wird diese qualitative Weiterentwicklung offensichtlich durch Erfahrung und Praxis. Kolodner (1983) zufolge verfügen Experten über bessere *episodische Definitionen* als Novizen. Episodische Definitionen bezeichnen dabei Erfahrungen über die Nutzungsmöglichkeiten ihres Wissens. Diese episodischen Definitionen verbessern und verfeinern sich demnach im Laufe des Expertiserwerbs durch Praxis. Nach Reimann und Chi (1989) ist diese Entwicklung darauf zurückzuführen, dass durch Erfahrung bestehende Wissensstrukturen immer wieder organisiert und reorganisiert werden. Mit dieser Vorstellung sind auch die *Schematheorien* (Minsky, 1975; Schank & Abelson, 1977) und insbesondere der *Skriptansatz* verwoben (Schank, 1982). Als Schemata werden Gedächtniseinheiten bezeichnet, die all das Wissen, über das man bezüglich von Objekten oder Situationen verfügt, zusammenfassen. Infolge von Erfahrungen können diese Schemata verändert werden.

Die vorgestellten Ansätze zur Wissensorganisation und -repräsentation verdeutlichen, dass Erfahrung und Übung in einer Domäne die Aneignung von Expertise bedingen. Kolodners Ansatz lässt bereits vermuten, dass die Art der Erfahrungen beim Expertiseerwerb eine entscheidende Rolle spielen könnte. Die folgenden Ansätze untersuchen die Art der Erfahrungen, die zum Erwerb von Expertise beitragen, sowie einzelne Schritte die dazu notwendig sind.

2.1.3.2 Die *deliberate-practice*-Theorie

Das Ergebnis einer von Ericsson et al. (1993) durchgeführten Literaturanalyse zum Erwerb von Expertise in unterschiedlichen Bereichen war, dass Übung zwar notwendig zur Ausbildung von Expertise ist, dass es aber von der Art der Übung abhängt, inwiefern die Leistungen dadurch maximiert werden. Personen müssen demnach motiviert sein, sich mit dem Lerngegenstand zu beschäftigen und sich in einer Aufgabe zu verbessern. Außerdem sollte zu Beginn des Lernens eine kurze Phase einer an das Vorwissen angepassten Instruktion stattfinden. Dabei sollte dem Lerner unmittelbares Feedback bezüglich seiner Leistung gegeben werden. Im Sinne der aus diesen Erkenntnissen entstandenen *deliberate-practice*-Theorie von Ericsson et al. (1993) entwickelt man Expertise in einer Domäne, in dem man zwischen Phasen der geleiteten Instruktion individuelle, wohldurchdachte und gezielte Übungen durchführt. Geleitete Instruktion bedeutet dabei, dass die Spieler bei der Übung durch einen Trainer angeleitet werden, während sie bei der wohldurchdachten Übung selbstständig konzentriert trainieren. Charness, Tuffiash, Krampe, Reingold und Vasyukovaa (2005) konnten in einer Studie mit Schachspielern verschiedener Nationalitäten zeigen, dass die Durchführung von individuellen wohldurchdachten Übungen (*deliberate practice*) essentiell für die Ausbildung von Expertise war. Die Leistung der Schachspieler in Schachturnieren korrelierte positiv mit der selbsteingeschätzten Zeit, die sie mit eigenständigem Üben, der Teilnahme an Turnieren und gezielter Instruktion zugebracht hatten. Die Zeit, die mit eigenständigem Üben verbracht wurde, erwies sich dabei als stärkster Prädiktor für die Turnierleistungen. Die Überlegenheit dieser Methode gegenüber der Teilnahme an Turnieren erklären die Autoren damit, dass bei *deliberate practice* das Übungsmaterial adaptiv an das Leistungsniveau der Person angepasst ist, während Schachspieler bei Turnieren in der Regel häufig gegen Spieler spielen, deren Leistungsniveau sich entweder weit über oder deutlich unter ihrem eigenen befindet. Aus dieser Annahme erschließt sich, dass eine Voraussetzung für das eigenständige gezielte Üben die Ausbildung metakognitiver Strategien zu sein scheint, die es erlauben, die eigene Leistung und den

eigenen Lernfortschritt zu überwachen, zu kontrollieren und zu entscheiden, welche Art des Trainings dem derzeitigen Niveau angepasst ist.

Wie wohldurchdachte Übung mit Expertise zusammenhängt, wird nach Ericsson et al. (1993) mit der sogenannten „monotonic benefits assumption“ (S. 368) verdeutlicht. Diese Annahme besagt, dass die Zeitmenge, die eine Person mit wohldurchdachter Übung verbringt, in einer monoton steigenden Beziehung zu ihrer Leistung steht. Durch *deliberate practice* steigt demnach das Experteniveau kontinuierlich an. Es gibt jedoch nach Ericsson et al. (1993) Anforderungen, die das wohldurchdachte Üben über längere Zeit stellt:

1. Ressourcen (*the resource constraint*): Zum Üben wird ausreichend Zeit, Energie, Lehrpersonal und Übungsmaterial gebraucht. Besonders bei Kindern ist die Unterstützung der Bezugspersonen wichtig, die dem Kind den Zugang zu den erforderlichen Ressourcen erst ermöglichen.
2. Motivation (*the motivational constraint*): Selbst wenn das Individuum in ausreichendem Maße über Ressourcen verfügt, bedeutet dies nicht zwangsmäßig, dass es diese auch zur Übung einsetzt. Wenn das Lernmaterial an sich nicht schon ansprechend genug ist, dient dem Individuum die Steigerung seiner eigenen Leistung als Ansporn. Bei Kindern greifen oft die Eltern ein und unterstützen ihre Kinder motivational.
3. Anstrengung (*the effort constraint*): In der Regel ist Üben kognitiv und/oder körperlich sehr anstrengend. Erschöpfung stellt sich ein, wenn über längere Zeit über ein gewisses Pensum hinaus gezielt trainiert wird. Dabei ist zwischenzeitliche Erholung wichtig, um gerade bei Anfängern eine Überforderung zu vermeiden.

Ericsson et al. (1993) zeigten, dass der Umfang an eigenständigem Üben mit steigendem Alter der Experten ansteigt. Die Autoren konnten in zwei Studien mit Musikern (Violinisten und Pianisten) Belege dafür sammeln, dass Unterschiede im Leistungsniveau, das die Personen im Erwachsenenalter erreicht hatten, auf die Menge an individueller Übung zurückgeführt werden konnte.

Deliberate practice scheint nicht nur für den Erwerb von Expertise relevant zu sein, sondern auch für die Aufrechterhaltung eines hohen Leistungsniveaus. Krampe und Charness (2006) konnten zeigen, dass wenn ältere Experten schlechtere Leistungen zeigten als sie in

jüngeren Jahren erbracht hatten, dies darauf zurückzuführen war, dass sie weniger Zeit mit gezielter Übung verbrachten als an ihrem persönlichen Leistungshöhepunkt.

An der *deliberate-practice*-Theorie kann kritisiert werden, dass sie keine Erklärungen dafür liefert, dass Expertise auch in Domänen entstehen kann, in denen direktes Feedback selten verfügbar ist. Zudem werden keinerlei Annahmen über die Veränderung kognitiver Funktionen geliefert. Nicht erklärbar ist nach Ericsson und Ward (2007) auch das Erleben des sogenannten Flow-Gefühls, das besonders Experten zuweilen erleben, wenn sie vollkommen vertieft in ihre aktuelle Tätigkeit sind (Csíkszentmihályi, 1990).

2.1.3.3 Das Fünf-Stufen-Modell des Fertigkeitserwerbs

Dreyfus und Dreyfus (1980) haben ein Modell entwickelt, das die Entwicklung vom Novizen zum Experten in fünf Stufen beschreibt. Das Modell bezieht sich dabei auf eine Form des Fertigkeitserwerbs, die sowohl gezielte Instruktion als auch Erfahrung einschließt. Die Autoren entwickelten ihr Modell anhand von deskriptiven Selbstbeschreibungen von Personen, die sich in unterschiedlichen Phasen des Expertiseerwerbes beim Erlernen einer Fremdsprache, dem Schachspielen und dem Fliegen befanden. Die Autoren argumentieren, dass lediglich basale Fertigkeiten auf abstrakten Regeln beruhen, während Fertigkeiten auf höheren Levels nur durch konkrete Erfahrungen ausgebildet werden.

Dreyfus und Dreyfus (1980) bezeichnen die erste Stufe in ihrem Modell des Fertigkeitserwerbs mit dem Begriff *Novice*. Auf dieser Stufe werden den Beginnern im Unterricht Begriffe und Regeln ohne situativen Kontext beigebracht. Z.B. können Flugschüler in dieser Phase Cockpitinstrumente lesen und bedienen. Die zweite Stufe wird *Competence* genannt und bezeichnet ein Stadium innerhalb des Fertigkeitserwerbes, das nur durch Erfahrung in der Bewältigung von realen Situationen oder durch wiederholte Hinweise durch den Übungsleiters auf bedeutsame Gegebenheiten erreicht werden kann. Verständliche Beispiele helfen dabei, Prinzipien (*guidelines*) für den Einsatz der erlernten Regeln zum Umgang mit bestimmten Konstellationen zu entwickeln. Auf die dritte Stufe, die als *Proficiency* bezeichnet wird, gelangt ein Lerner dadurch, dass er mit der Zeit immer mehr unterschiedlichen Situationen ausgesetzt ist, die er bewältigen muss. Er bildet sogenannte *Maxime* aus. Maxime sind Prinzipien zur Bewältigung eines Problems, das in einem größeren Kontext wahrgenommen wird. Jede Aktion ist somit auf ein höheres Ziel ausgerichtet. So bedenkt ein Schachspieler auf dieser Stufe bei jedem einzelnen Zug dessen Konsequenzen für ein strategisches Ziel (z.B. Angriff).

Solange sich Personen auf den ersten drei Stufen des Fertigkeitserwerbes befinden, brauchen sie Regeln, Prinzipien und Maxime und müssen diese explizit erinnern, um eine Verbindung zwischen der aktuellen Situation und der generellen Situation bzw. übergreifenden Zielen zu schaffen. Auf der vierten Stufe, die *Expertise* genannt wird, hat sich ein Wandel von analytischem Vorgehen zu intuitivem Handeln vollzogen. Es werden keine expliziten Regeln mehr gebraucht, da die Person intuitiv das in der Situation angemessene Verhalten wählt. Durch die ausgeprägte Erfahrung wurde jede Situation mit einer spezifischen Verhaltensantwort assoziiert.

Der Fertigkeitserwerb ist mit der Stufe der Expertise abgeschlossen. Die fünfte Stufe des Modells, die *Mastery* genannt wird, kann jedoch von Experten erlebt werden, wenn sie ihre Arbeit zeitweise vollkommen ohne bewusste Aufmerksamkeit verrichten können. Ein Beispiel dafür ist der geübte Autofahrer, der sich im Nachhinein an bestimmte Streckenabschnitte seiner Fahrt nicht einmal mehr erinnern kann.

Die Entwicklung vom Novizen zum Experten bzw. Master kann laut Dreyfus und Dreyfus (1980) mit der Transformation der vier mentalen Funktionen Erinnerung, Erkennung, Entscheidung und Aufmerksamkeit erklärt werden. In Tabelle 2.1 ist die Transformation der mentalen Funktionen über die verschiedenen Stufen des Fertigkeitserwerbs hinweg dargestellt.

Tabelle 2.1

Transformation mentaler Funktionen auf den Stufen des Modells des Fertigkeitserwerbs nach Dreyfus und Dreyfus (1980)

Stufe Funktion	Novice (1)	Competence (2)	Proficiency (3)	Expertise (4)	Mastery (5)
Erinnerung	nicht situativ	situativ	situativ	situativ	situativ
Erkennung	diskret	diskret	holistisch	holistisch	holistisch
Entscheidung	analytisch	analytisch	analytisch	intuitiv	intuitiv
Aufmerksamkeit	überwachend	überwachend	überwachend	überwachend	absorbiert

Auf der Stufe des Novizen befinden sich sämtliche mentale Funktionen noch in ihrer primitiven Form. Auf Stufe zwei verfügen Personen über situative Erinnerungen. Ähnlichkeiten von Situationen werden erkannt, was erst durch die zunehmende Erfahrung mit dem jeweiligen Inhaltsbereich möglich ist. Die Erkennung von situativen Komponenten wird

erst ganzheitlich (holistisch), wenn Ähnlichkeiten bei ganzen Situationen und nicht nur einzelnen Komponenten einer Situation wahrgenommen werden (Stufe 3). Dabei wird auch erlernt, das Wichtige an einer Situation zu erkennen. Entscheidungen können erst auf Stufe vier intuitiv ohne bewusste Verarbeitung erfolgen. Auf der letzten Stufe kann sogar die Überwachung und Evaluation des Verhaltens aufgegeben werden.

2.1.3.4 Die cognitive-flexibility-Theorie

Die *kognitive Flexibilitätstheorie (cognitive-flexibility-Theorie)* kann insbesondere auf den vorangeschrittenen Wissenserwerb und die Anwendung von Wissen in komplexen, schlecht strukturierten Domänen angewandt werden (Spiro, Coulson, Feltovich & Anderson, 1988). Spiro and Jehng (1990) definieren kognitive Flexibilität als „the ability to adaptively reassemble diverse elements of knowledge to fit the particular needs of a given understanding or problem-solving situation” (S. 169). Die Autoren beziehen sich mit ihrer Theorie nicht auf das anfängliche Lernen in einer Domäne sondern auf das Lernen, das nach einer einleitenden Lernphase stattfindet. Dabei wird angenommen, dass die Lernenden dann nicht mehr als absolute Novizen, sondern als Fortgeschrittene (*intermediates*) zu bezeichnen sind. Diese Fortgeschrittenen brauchen nach Ansicht der Autoren bestimmte Lernvoraussetzungen, um sich weiter verbessern zu können und den Status des Experten zu erreichen. Ein Kennzeichen von Experten ist deren flexibler Umgang mit Problemsituationen (Gruber & Mandl, 1995). Dies ist der Theorie der kognitiven Flexibilität zufolge darin begründet, dass Experten besonders reichhaltige Erfahrungen in einer Domäne gemacht haben, was zu einerseits umfangreichem aber andererseits auch differenziertem und facettenreichem Wissen geführt hat (Spiro, Feltovich, Jacobson & Coulson, 1991; Spiro & Jehng, 1990). Aus der Annahme, dass sich die Fähigkeit zum flexiblen, angepassten Umgang mit Problemsituationen als Folge von unterschiedlichen Erfahrungen mit dem Lerngegenstand entwickelt, wird geschlossen, dass zum Aufbau von Expertise eine lineare Wissensaufnahme (wie z.B. aus einem Buch) kontraproduktiv ist. Spiro et al. (1991) führen daher verschiedene Prinzipien an, die das optimale Lernen gewährleisten sollten. Es werden *flexible Lernumgebungen* gefordert, die die Möglichkeit bieten, identische Wissensinhalte in *unterschiedlichen Kontexten* und aus *unterschiedlichen Perspektiven* darzustellen und zu erlernen. Eine zu starke Reduktion der Komplexität erachten die Autoren als kontraproduktiv für das fortgeschrittene Lernen. Sie empfehlen das Lernen *anhand von Fallbeispielen* (z.B. in der Medizin). Außerdem sollte das zu erlernende Wissen nicht in kleinere Sektionen unterteilt werden. Diese Unterteilung wird der *Komplexität des Lerninhaltes* nicht gerecht. Aus dieser Theorie lassen sich Implikationen

für das computerbasierte Lernen z.B. mithilfe von Hypertext ableiten (Jacobson & Spiro, 1995). Die Theorie der kognitiven Flexibilisierung fokussiert besonders auf ein Merkmal von Experten in komplexen Domänen: ihrer Fähigkeit zum schnellen und flexiblen Problemlösen. Aus dieser Theorie lässt sich ableiten, dass reichhaltige Erfahrungen notwendig sind, damit Personen Expertise in diesen komplexen Domänen erwerben können.

In diesem Kapitel wurden verschiedene Theorien zur Entwicklung von Expertise vorgestellt. Es hat sich dabei gezeigt, dass zwischen diesen Theorien ein weitestgehender Konsens darüber besteht, dass domänenspezifische Erfahrungen und Übung eine bedeutende Rolle beim Expertiseerwerb spielen. Während man sich in grundlegenden Theorien eher auf die effiziente Speicherung und Organisation deklarativer Wissensinhalte bezieht (*Chunking-Theorie*, *Pattern-Recognition-Ansatz* und *Skilled-Memory-Theorie*), wird in weiterführenden Theorien versucht den Aufbau prozeduralen Wissens und automatisierter, intuitiver Vorgänge zu beschreiben (ACT*-Modell) sowie eine Spezifikation der Merkmale der zum Expertiseaufbau erforderlichen Erfahrungen vorzunehmen (*deliberate-practice-Theorie* und *cognitive-flexibility-Theorie*). Dabei kann nicht eine einzige Theorie als angemessen für jegliche Aspekte des Expertiseerwerbs in einer Domäne betrachtet werden. Um Vorhersagen über den Expertiseerwerb in den einzelnen Teilaspekten einer komplexen Domäne wie dem Autofahren zu treffen, können unterschiedliche Theorien herangezogen werden. In Kapitel 2.3.5 wird darauf eingegangen, inwieweit die verschiedenen Theorien des Expertiseerwerbs angemessen zur Vorhersage der Leistungsentwicklung in verschiedenen Teilbereichen der Domäne Autofahren erscheinen.

2.1.4 Methoden der Expertiseforschung

Möchte man Vorhersagen, die man für eine Domäne bzgl. des Expertenwissens und -könnens sowie bzgl. der Ausbildung von Expertise getroffen hat, überprüfen, benötigt man geeignete Forschungsdesigns. Zur gründlichen Erforschung von Expertise in einer Domäne schlagen Ericsson und Smith (1991a) ein System vor, das der Erforschung von Expertise in einem Inhaltsbereich einen Rahmen geben soll. Beim so genannten *expert performance approach* können drei wichtige Stationen im Prozess der empirischen Analyse von Expertise unterschieden werden. Das Ziel der ersten Stufe in diesem Prozess stellt die Operationalisierung von Expertise in der ausgewählten Domäne dar. Dazu müssen repräsentative Aufgaben für Labor und Feldstudien erstellt und empirisch getestet werden. Auf der zweiten Stufe werden die der Expertenleistung zugrunde liegenden Mechanismen

genauer betrachtet. Wurden im ersten Schritt also Aufgaben identifiziert, in denen Experten im Gegensatz zu Novizen sehr gute Leistungen erbringen, wird im zweiten Schritt zu ermitteln versucht, welche kognitiven oder motorischen Prozesse zu den beobachteten Leistungsunterschieden beigetragen haben. Dazu werden in der Regel Messmethoden benötigt, die die Leistung auf der Prozessebene erfassen. Zu solchen Methoden gehören beispielsweise die Blickbewegungsmessung und das *Laute Denken*. In der dritten Phase der Expertiseforschung soll ermittelt werden, wie der Expertiseerwerb in der betrachteten Domäne erfolgt. Dazu ist es wichtig zu erfassen, wie und unter welchen Umständen die verschiedenen Merkmale der Expertise sich entwickeln. Für jedes der drei Stadien gibt es jeweils unterschiedliche experimentelle Vorgehensweisen, die für die Fragestellung besonders geeignet sind. Die Entwicklung von repräsentativen Aufgaben und die Identifikation der Leistung zugrunde liegender Mechanismen erfolgt in der Regel in Experten-Novizen-Vergleichen. Die Expertiseentwicklung hingegen lässt sich am zuverlässigsten in Trainings- oder Längsschnittstudien untersuchen. Im Folgenden werden die verschiedenen Methoden der Expertiseforschung dargestellt und ihre jeweiligen Vor- und Nachteile erörtert.

2.1.4.1 Experten-Novizen-Vergleiche

Bei den meisten Studien zur Erforschung von Expertise handelt es sich um Querschnittstudien. Dabei wird die Leistung von Novizen in domänenspezifischen Aufgabenstellungen mit der Leistung von Experten verglichen. Das Novizen-Experten-Paradigma dient dazu, interindividuelle Unterschiede durch zwei extreme Leistungsgruppen, die zuvor als solche bekannt sind, zu identifizieren. Die Zuteilung zu Versuchsgruppen ist sehr eindeutig und valide. Das Novizen-Experten-Paradigma erlaubt Studien, die über das Untersuchungslabor hinaus valide sind, denn es können auch Personen einbezogen werden, die sich in Domänen des alltäglichen Lebens als Experten erwiesen haben. Experten-Novizen-Vergleiche zeigen auf, in welchen Merkmalen sich zwei Leistungsgruppen unterscheiden und können daher als Grundlage für weitere Forschung z.B. zum Erwerb von Expertise dienen. Das Paradigma ist besonders wertvoll als Validierungsmethode zur Beurteilung von, für die jeweilige Domäne repräsentativen, Aufgaben. In Leistungstests zu einem bestimmten Inhaltsbereich sollten die Experten in dieser Domäne den Novizen deutlich überlegen sein. Dies kann als Hinweis darauf betrachtet werden, dass einzelne Aufgaben bzw. ganze Testverfahren kriteriumsvalide dasjenige Merkmal messen, zu deren Messung sie konstruiert wurden.

Nach Gruber (1994) ist dieser *kontrastive* Ansatz als Forschungsmethode allerdings nicht völlig unumstritten. Problematisch bei diesem Vorgehen ist, dass sich Gruppen von Personen mit unterschiedlichem Expertisegrad in der Regel nicht nur in der Ausprägung der Expertise in einer Domäne unterscheiden, sondern auch durch viele andere Merkmale wie z.B. dem Alter und dem sozio-ökonomischen Status (Reimann & Rapp, 2008). Zudem kann durch die querschnittliche Vorgehensweise beim Experten-Novizen-Vergleich nur bedingt untersucht werden, wie sich Expertise entwickelt und welche Faktoren dabei eine Rolle spielen. Was auf den Zwischenstufen passiert, die der Novize erreicht, bis er zum Experten wird, bleibt weitestgehend unklar (vgl. van Gog, Ericsson, Rikers & Paas, 2005). Experimentelle und längsschnittliche Ansätze sind demgegenüber dazu geeignet, Kausalitäten zu identifizieren. Zur Untersuchung der Expertiseentwicklung muss daher in der Regel auf zusätzliche Methoden zurückgegriffen werden.

2.1.4.2 Längsschnittstudien

Bei Längsschnittuntersuchungen können Personen auf ihrem Weg vom Novizen zum Experten zu mehreren Messzeitpunkten untersucht werden (z.B. Gruber, Renkl & Schneider, 1994). Diese Vorgehensweise ist demnach dazu geeignet, den Expertiseerwerb zu untersuchen und stellt zudem die beste Möglichkeit dar, theoretische Modelle zur Expertiseentwicklung zu überprüfen. Da die Entwicklung von Expertise in der Regel jahrelange Beschäftigung mit einem Inhaltsbereich erfordert, sind Längsschnittuntersuchungen in der Expertiseforschung eher selten. Bei den wenigen Studien, die längsschnittlich durchgeführt werden, handelt es sich zudem meist um Fallstudien. Das heißt, es wird nur eine einzige Person über den ganzen Zeitraum hinweg beobachtet. Dies hat den Nachteil, dass Erkenntnisse, die aus solchen Studien stammen, in der Regel nicht als generalisierbar gelten können. Als vorteilhaft bei Fallstudien kann jedoch gesehen werden, dass sie zumeist in einem natürlichen Setting durchgeführt werden. Das heißt die Person, die sich zum Experten entwickelt, befindet sich dabei in ihrem natürlichen Umfeld. Sie sucht sich selbst aus, wann, wo und wie viel sie in ihrer Domäne lernt. Dieser Umstand birgt allerdings einen weiteren Nachteil für die Forschung, denn der Expertiseerwerb ist in hohem Maße von der Motivation einer Person abhängig. Somit sind Längsschnittstudien schlecht planbar, da man, ohne auf ein naturalistisches Setting zu verzichten, den Expertiseerwerb von außen kaum beeinflussen kann. Mit hohen Ausfallquoten muss zudem gerechnet werden, da nicht alle Personen, die beginnen auf einem Gebiet Wissen zu sammeln und Fähigkeiten auszubilden, zwangsläufig irgendwann zum Experten auf diesem Gebiet werden.

Aus den oben genannten Gründen werden selten Ergebnisse aus Längsschnittuntersuchungen zum Expertiseerwerb berichtet. Charness (1989) berichtet von einer Längsschnittstudie in der Domäne Schach, die jedoch ursprünglich nicht als Längsschnittstudie geplant war. Dabei traf Charness zufällig auf eine Versuchsperson aus einer früheren Studie, die sich in den zwischenzeitlich neun vergangenen Jahren vom durchschnittlich leistungsstarken Turnierspieler zum Schachmeister auf internationalem Niveau entwickelt hatte. Diese Person war nun bereit, ein zweites Mal an einer Studie teilzunehmen. Bei der Interpretation der Ergebnisse dieser Studie lag jedoch die Einschränkung vor, dass die Untersuchung nicht die komplette Entwicklung vom Novizen zum Experten verfolgen konnte, da sich die Person bereits zum Zeitpunkt der ersten Erhebung auf einer Stufe zwischen Novize und Experte befand.

In einer anderen Längsschnittuntersuchung konnten Gruber, Renkl und Schneider (1994) belegen, dass das domänenspezifische Gedächtnis sich im Laufe des Expertiseerwerbs tatsächlich verbessert. Dazu wurden 27 Schachexperten und -novizen, die an einer Studie drei Jahre zuvor teilgenommen hatten (Opwis, Gold, Gruber & Schneider, 1990), zu einer weiteren Messung eingeladen. Ein Ergebnis dieser Untersuchung war, dass die schachspezifischen Gedächtnisleistungen über die Zeit angestiegen waren. Es zeigte sich zudem, dass diejenigen Personen, die sich innerhalb der drei vergangenen Jahre als Aussteiger aus der Expertisekarriere erwiesen hatten, schon zum ersten Messzeitpunkt schwächere schachspezifische Merkleistungen gezeigt hatten als diejenigen Experten, die ihre Karriere weiter verfolgt hatten. Dieses Ergebnis weist darauf hin, dass einmalige Experten-Novizen-Vergleiche nur bedingt valide sein können, wenn sich unter den Experten Personen befinden, die später aussteigen werden. Zur Erfassung der dafür verantwortlichen spezifischen Merkmale sind Längsschnittuntersuchungen unabdingbar.

2.1.4.3 Trainingsstudien

Eine Alternative zur Längsschnittstudie ist die Trainingsstudie, bei der die Leistung der Probanden vor und nach einem spezifischen Training erfasst wird. Es handelt sich dabei zwar auch um ein längsschnittliches Untersuchungsdesign, das aber im Gegensatz zu einer echten Längsschnittstudie nicht den „natürlichen“ Gang der Entwicklung betrachtet. In der Regel erfolgt zunächst eine Vorher-Messung, die Aufschluss darüber geben soll, auf welchem Leistungsstand sich die Teilnehmer befinden. Danach erfolgt eine Intervention, also ein Training, bei dem sich die Teilnehmer intensiv mit dem Untersuchungsgegenstand beschäftigen. Dabei herrschen in der Regel Vorgaben bezüglich Lernzeit, Lerninhalten und

der Abfolge dieser Lerninhalte. Nach der Trainingsphase erfolgt eine Nachher-Messung, die dazu dient den Trainingseffekt festzustellen. In vielen Trainingsstudien wird noch eine Follow-Up-Messung nach einem bestimmten Zeitintervall ohne zwischenzeitliche Intervention durchgeführt. Dadurch soll überprüft werden, inwieweit Trainingseffekte nachhaltig fortbestehen. Die Trainingsstudie eignet sich aus verschiedenen Gründen nur begrenzt zur Untersuchung der Entwicklung von Expertise. Zum einen erfordert ein Training, das die Teilnehmer vom Novizen zum Experten bringen soll, enormen Aufwand von den Teilnehmern, da sie in relativ kurzer Zeit eine Entwicklung durchlaufen sollen, die unter natürlichen Bedingungen mehrere Jahre erfordert. Zum anderen stellt es sich als problematisch dar, dass Trainingsprogramme in der Regel nach den bereits bestehenden Annahmen zur Entwicklung der Expertise gestaltet werden. Dieses methodische Problem bewirkt, dass die Richtigkeit bestehender Vorstellungen zum Expertiseerwerb weder widerlegt noch bestätigt werden kann. Aus diesen Gründen eignen sich Trainingsstudien gut dazu, die Wirksamkeit einer Interventionsmaßnahme zu überprüfen. Zur Untersuchung der Expertiseentwicklung sind ihre Möglichkeiten eher als begrenzt zu betrachten.

2.1.5 Besonderheiten der Untersuchung von Expertise im Autofahren

Im Folgenden wird dargestellt, inwieweit sich die Konzepte der Expertiseforschung auf die Domäne Autofahren anwenden lassen. Zunächst ist die Definition der Expertise bzw. des Experten zu betrachten. Als grundlegend wird zunächst die allgemein formulierte Definition von Posner (1988) betrachtet, die besagt, dass ein Experte in seiner Domäne dauerhaft sehr hohe Leistungen erbringt. Übertragen auf das Autofahren bedeutet dies, dass ein Experte vorwiegend sehr gut fährt. Die Leistung von PKW-Fahrern wird in der Regel nach dem Bestehen der Fahrerlaubnisprüfung nicht mehr gemessen. Die Leistungsmessung geht, selbst in der Fahrerlaubnisprüfung, über eine Klassifikation in „gut genug zum Bestehen“ und „nicht gut genug zum Bestehen“ nicht hinaus. Lediglich bei Autorennen, wie z.B. in der Formel 1 oder bei der Deutschen Tourenwagen Meisterschaft (DTM), wird die Leistung von Fahrern gemessen. Als Kriterium hierfür wird die Geschwindigkeit, mit der der Fahrer in seinem Wagen eine vorgegebene Strecke auf einem Parcours absolviert, herangezogen. Die Fahrer werden anhand dieser Zeiten in eine Reihenfolge gebracht. Gutes Fahren ist hierbei demnach besonders schnelles Fahren bei gleichzeitig hoher Kontrolle über das Fahrzeug. Als Kriterium für überdauerndes sehr gutes Fahren im Realverkehr kann dieses Kriterium jedoch nicht dienen. Dies bestätigte eine frühe Studie von Williams und O’Neill (1974), die zeigte, dass Rennfahrer im Realverkehr sogar in mehr Unfälle verwickelt waren als Personen mit ähnlich

hoher Fahrerfahrung, bei denen es sich nicht um Rennfahrer handelte. Rennfahrer können demnach nicht unbedingt als Experten für das Fahren im Straßenverkehr gelten.

Als Kriterium für Fahrexpertise sollte sich besser die Anzahl der von einer Person verursachten Unfällen und Verkehrsdelikte über einen bestimmten Zeitraum eignen. Je weniger Unfälle eine Person in diesem Zeitraum verursacht und je weniger Verkehrsdelikte sie begangen hat, desto höher ist die Expertise einzuschätzen. Wichtig ist dabei allerdings, dass die Anzahl der Unfälle und Delikte an der Fahrleistung relativiert werden muss. Expertise entsteht durch Erfahrung und Übung. Dahingehend sollte die Domäne des Autofahrens keine Ausnahme darstellen. Je mehr Kilometer man mit dem Auto fährt, desto besser sollte man als Autofahrer werden. Um die Domäne des PKW-Fahrens für die Methoden der Expertiseforschung zugänglich zu machen, muss ein eindeutiges Außenkriterium für Fahrexpertise festgesetzt werden. Wie in Kapitel 2.1.4.1 erwähnt, werden in der Expertiseforschung zumeist kontrastive Studien eingesetzt. Zur Durchführung muss entschieden werden, wann eine Person als Experte zu bezeichnen ist und auf wen die Bezeichnung Novize zutrifft.

In der Expertiseforschung wird zwischen *wohl definierten* und *schlecht definierten* Domänen unterschieden (Gruber & Mandl, 1996). Dabei sind wohl definierte Domänen dadurch gekennzeichnet, dass es eindeutige Problemstellungen gibt, die richtigen Lösungswege bekannt sind und Kriterien für Erfolg mühelos identifiziert werden können. Schlecht definierte Domänen zeichnen sich demgegenüber durch unklare Problemstellungen, fehlende festgelegte Regeln zur Bewertung des Erfolges und fehlende eindeutige Lösungsstrategien aus. Das Autofahren ist – wie die meisten *lebensweltlichen* Domänen – als eine eher schlecht definierte und komplexe Domäne zu bezeichnen (Sturzbecher, 2008). Im Gegensatz zu beispielsweise dem Schachspiel, das besonders häufig als Gegenstand der Expertiseforschung untersucht wurde, fällt es beim Autofahren schwer, Personen ein objektivierbares Leistungsniveau zuzuordnen. Zudem liegt ein fester Wissensbestand, analog zu den unveränderlichen Regeln des Schachspiels, in der Domäne des Autofahrens in geringerem Maße vor. Es existieren zwar Verkehrsregeln, die allerdings geändert werden können. Auch in der Fahrschule erlernte Faustregeln, z.B. zum Bremsweg, sind veränderlich, da sie von technischen Fortschritten der Fahrzeuge abhängig sind. Zudem ist nicht klar, inwieweit erfahrene Fahrer über Kenntnisse bzgl. dieser Verkehrs- und Faustregeln verfügen und ob diese überhaupt das sichere Fahren beeinflussen. Außerdem stellt das Autofahren neben kognitiven auch psychomotorische Anforderungen an den Fahrer. Darüber hinaus

werden Fähigkeiten benötigt, die eine Anpassung an durch ständige Veränderungen gekennzeichnete Umwelanforderungen ermöglichen.

Das Autofahren gleicht damit eher gering strukturierten Domänen, wie beispielsweise diversen Sportarten. Die Domäne Sport im Allgemeinen hat sich in der Vergangenheit als ein besonders ertragreicher Inhaltsbereich zur Validierung von Expertisekonzepten, die ursprünglich in anderen Domänen entwickelt wurden, erwiesen (Williams, A. M. & Ericsson, 2005). Gruber und Lehmann (2007) führen einige Besonderheiten sportlicher und musischer Domänen auf, deren Ähnlichkeit zur Domäne Autofahren im Folgenden überprüft wird.

1. Es bestehen Anforderungen an die Psychomotorik einer Person, denn koordinierte, teilweise komplexe (Körper-)Bewegungen, zuweilen im Zusammenspiel mit anderen Personen, werden verlangt.
2. Es bestehen Zeiteinschränkungen oder -vorgaben, z.B. erfordern Prozesse schnelle Reaktionen.
3. Das effektive Zusammenspiel unterschiedlicher Fähigkeiten ist notwendig, wobei die Schwierigkeit darin besteht, dass sich einzelne Fähigkeiten in unterschiedlicher Geschwindigkeit entwickeln können.
4. Wenn am Gesamtsystem mehrere Personen beteiligt sind, können diese unterschiedliche Rollen, die verschiedenartige Anforderungen stellen, einnehmen (z.B. Angriff- und Abwehrspieler beim Handball).

Insgesamt finden die meisten Sportarten in einer dynamischen, sich stets wandelnden Umgebung statt. Dabei handelt es sich in der Regel um extreme Stresssituationen, deren Anforderungen den Sportlern Leistungen bis zu ihren physischen und kognitiven Belastbarkeitsgrenzen abverlangen (Williams, A. M. & Ericsson, 2005). Punkt 1 kann als Besonderheit für die Domäne des Autofahrens übernommen werden. Beim Autofahren werden komplexe Bewegungsabläufe (z.B. beim Schalten, Abbiegen usw.) relevant, die aus verschiedenen Teilschritten bestehen, die koordiniert werden müssen. Diese Abläufe sind im Gegensatz zum Mannschaftssport nicht über mehrere Personen verteilt. Wenn man aber den Verkehrskontext mit den verschiedenen Verkehrsteilnehmern betrachtet, wird die Notwendigkeit der Koordination deutlich (z.B. an Kreuzungen oder beim Auffahren auf die Autobahn). Analog zu Punkt 2 existieren beim Autofahren Zeiteinschränkungen. Schnelle

Reaktionen sind notwendig, wenn andere Verkehrsteilnehmer oder das eigene Fahrzeug sich unvorhersehbar verhalten. Auch Wettergegebenheiten und Streckeneigenschaften beeinflussen das Verhalten des Fahrers. Die in Punkt 3 angeführte Besonderheit in sportlichen Domänen lässt sich ebenfalls auf das Autofahren übertragen. Besonders zu Beginn ihrer Fahrkarriere sollten Fahranfänger über umfangreiches deklaratives verkehrsbezogenes Wissen verfügen, während andere Fertigkeiten, wie beispielsweise die Fahrzeugbeherrschung oder das schnelle Überblicken komplexer Verkehrssituationen, möglicherweise noch nicht optimal ausgebildet sind. Analog zu Punkt 4 nehmen verschiedene Verkehrsteilnehmer unterschiedliche Rollen im Verkehrsgeschehen ein. So muss man sich als Fahrer z.B. in Fußgänger, die am Fahrbandrand stehen, hinein versetzen, um antizipieren zu können, ob sie sich in gefährdender Weise verhalten werden. Ein anderes Beispiel für das Zusammenspiel verschiedener Rollen im Verkehr stellen Überholvorgänge dar. Der Überholende muss dabei erkennen, wann er sich im toten Winkel des Überholten befindet und besonders in dieser Phase aufmerksam das Verhalten des anderen beobachten. Dieser muss sein Fahrverhalten seiner Rolle anpassen, um einer Gefährdung vorzubeugen (z.B. seine Geschwindigkeit nicht erhöhen).

Entscheidungen im Sport zu treffen erfordert vom Handelnden in effizienter Weise Entscheidungen über sein zukünftiges Agieren und Reagieren zu fällen. Expertise im Sport drückt sich generell darin aus, dass die Experten akkuratere und schnellere Entscheidungen treffen können als Novizen (Hodges, Huys & Starkes, 2007). Auch beim Autofahren müssen Entscheidungen bzgl. der Planung des eigenen Handelns getroffen werden, die an die jeweilige Situation angepasst werden müssen. Für diese Entscheidungen bleibt nicht selten wenig Zeit.

Die Domäne Autofahren erscheint in ihren Anforderungen demnach vergleichbar mit sportlichen Domänen zu sein. Ein empirisch gesichertes Kompetenzmodell für den Erwerb von Fahrexpertise liegt bisher noch nicht vor. Der Erwerb von Expertise in der Domäne Autofahren findet zunächst strukturiert und in einer zweiten Phase eher unstrukturiert statt. Während der Fahrausbildung gibt es zwar noch eine mehr oder weniger feste Vorgehensweise, in der der Stoff des Theorieunterrichts vermittelt wird und auch für den praktischen Unterricht gibt es bestimmte Regeln. Mit der formellen Fahrausbildung wird allerdings nur ein Grundstein für die Ausbildung von Expertise in dieser Domäne gelegt. Der Fahrschüler lernt alle Verkehrsregeln, das Bedienen des Fahrzeuges sowie physikalische Regeln, denen das Fahrzeug als bewegter Gegenstand unterliegt. Zusätzlich wird auch auf

Risiken von gefährlichen Verhaltensweisen im Verkehr hingewiesen. Der wirkliche Expertiseerwerb erfolgt jedoch erst in den ersten Jahren nach dem Erwerb der Fahrerlaubnis. Kurz nach der Fahrausbildung ist das Unfallrisiko der Führerscheinneulinge tatsächlich noch stark erhöht und verringert sich mit zunehmender selbstständig erworbener Fahrerfahrung deutlich (Schade, 2001). Das heißt, dass entscheidend für den Expertiseerwerb die Zeit des ersten eigenständigen Fahrens zu sein scheint. Die Situationen, in denen der Lernprozess dabei abläuft, sind unstrukturiert und informell (Grayson & Sexton, 2002). Es handelt sich dabei eher um *inzidentelles* oder beiläufiges Lernen. Junge Fahrer lernen demnach während des Fahrens, ohne sich dabei bewusst darüber zu sein, dass sie gerade lernen. Ein wichtiges Merkmal des inzidentellen Lernens ist die fehlende Intentionalität (Oerter, 1997). Dabei fehlen Absichten nicht komplett, sie sind lediglich nicht auf die Lernaktivität ausgerichtet. Somit verfolgen auch Fahranfänger beim Fahren nicht das Ziel ihre Fertigkeiten zu verbessern, sondern in der Regel, von einem Ort zu einem anderen zu gelangen. Als ein weiteres Merkmal des inzidentellen Lernens wird die fehlende Bewusstheit genannt. Der Lernerfolg stellt sich ein, ohne dass dem Lerner bewusst ist, dass oder wie er lernt. Als drittes Kriterium des inzidentellen Lernens führt Oerter (1997) die fehlende Aufmerksamkeitszuwendung auf. Damit ist allerdings nicht gemeint, dass beim beiläufigen Lernen keine Aufmerksamkeitsprozesse erfolgen. Lediglich wird die Aufmerksamkeit nicht selektiv auf die zu lernenden Prozesse gerichtet. Beim Autofahren wird in der Regel durchgehend Aufmerksamkeit benötigt, die aber nicht auf das Lernen ausgerichtet ist, sondern in der Regel auf den Verkehr außerhalb des Autos, durch den sich das Fahrzeug bewegt. Teilhandlungen laufen dabei allerdings mit steigender Erfahrung in einem immer größer werdenden Maße unbewusst ab. Zu diesen Teilhandlungen gehören z.B. das Schalten, das Spurhalten und motorische Handlungen beim Abbiegen. Beim inzidentellen Lernen können neben diesen prozeduralen Wissensinhalten auch deklarative Wissensinhalte erworben werden (Oerter, 1997). So kann ein junger Fahrer beim eigenständigen Fahren z.B. lernen, in welchen Gebieten, die er häufig passiert, besonders aufmerksam und vorsichtig gefahren werden muss, weil dort z.B. häufig Kinder spielen.

Durch die Einführung des optionalen Begleiteten Fahrens ab 17 Jahren (Stiensmeier-Pelster & Schöne, 2005) findet bei denjenigen Personen, die sich für dieses Modell entschieden haben, in der Regel auch in der Phase des ersten selbstständigen Fahrens noch explizites Lernen statt. Begleitpersonen, die im ersten Jahr nach dem Erwerb der Fahrerlaubnis an jeder Fahrt eines Fahranfängers teilhaben müssen, weisen den Fahranfänger

vermutlich auf besondere Gefahren hin, erinnern ihn an Verkehrsregeln und unterstützen ihn bei der Entscheidungsfindung in Verkehrssituationen, denen er im praktischen Fahrschulunterricht nicht begegnet ist. Dieses Lernen findet jedoch ebenfalls in unstrukturierter Form statt.

Zusammenfassend folgt beim Erwerb von Fahrexpertise auf die Phase intentionalen Lernens in der Fahrschule eine Phase inzidentellen Lernens beim eigenständigen Fahren im Realverkehr. Dabei wird Expertise in unterschiedlichen Teilaspekten vermutlich auf unterschiedliche Weise erworben.

2.1.6 Zusammenfassung und Forschungsdesiderate

Die Expertiseforschung beschäftigt sich mit den Merkmalen und der Entwicklung von Expertise. In der vorliegenden Arbeit werden Befunde der Expertiseforschung auf den Inhaltsbereich des Autofahrens übertragen. Obwohl in der Expertiseforschung in der Regel überdurchschnittliche Leistungen in bestimmten Domänen untersucht werden, scheint der Begriff der Expertise auch nützlich zum Verständnis der Entwicklung von Prozessen, Mechanismen und Kompetenzen, die in Verbindung mit alltäglichen Fähigkeiten stehen (Dreyfus & Dreyfus, 1986). Die Merkmale der Expertise haben sich in der Forschung über viele Domänen hinweg bestätigt. Die wichtigsten Befunde sind dabei, dass Experten Novizen in domänenspezifischen Aufgaben übertreffen, das heißt sie lösen diese schneller und machen weniger Fehler dabei. Expertise ist domänenspezifisch: Die Überlegenheit der Experten besteht nicht unbedingt auch in anderen Domänen und je spezifischer die vorgegebene Aufgabe für eine Domäne ist, desto deutlicher zeigt sich der Vorteil der Experten. Es ist daher zu erwarten, dass Expertise in der Domäne Autofahren ähnliche Merkmale aufweist.

Nicht so eindeutig zeigen sich die Annahmen und Befunde bzgl. der Entwicklung von Expertise. Es existieren unterschiedliche Modelle zum Expertiseerwerb. Die Passung der Modelle zur jeweiligen Domäne muss zunächst theoretisch begründet und anschließend empirisch überprüft werden. Soll dies für die Domäne Autofahren erfolgen, müssen die Besonderheiten dieser Domäne berücksichtigt werden. Der Expertiseerwerb erfolgt beim Autofahren ähnlich wie bei manchen Sportarten für die einzelnen Teilfertigkeiten auf unterschiedliche Weise. Deklaratives Verkehrswissen und auch grundlegende prozedurale Fahrfertigkeiten werden zunächst durch direkte Instruktion erworben. Danach erfolgt das Lernen von deklarativen und prozeduralen Lerninhalten eher inzidentell beim eigenständigen Fahren.

Inwiefern sich bestehende Modelle des Expertiseerwerbs auf die beginnende Entwicklung der Fahrexpertise bzw. ihre verschiedenen Teilaspekte auswirken, ist bisher noch nicht ausreichend untersucht. Dies soll in dieser Arbeit durch eine Längsschnittstudie erfolgen, bei der der Erwerb von Fahrexpertise im natürlichen Kontext untersucht wird. Um zu entscheiden, welche Kriterien in der Studie erhoben werden sollen, da sie für die Fahr- und Verkehrsexpertise relevant sind, muss zunächst aufgearbeitet werden, in welchen Aspekten sich gute Fahrer von schlechten Fahrern unterscheiden.

2.2 Das Risiko junger Fahranfänger

Im Jahr 2009 verunglückten in Deutschland 4152 Personen im Straßenverkehr. Das waren 7,3 % weniger als im Jahr 2008 (Statistisches Bundesamt, 2010b). Damit setzte sich der bereits seit 18 Jahren durchgängige Trend rückläufiger Zahlen im Verkehr tödlich verunglückter Personen fort. Laut Statistischem Bundesamt (2010b) kann man die Gruppe der 18- bis 24-Jährigen – die Gruppe der Fahranfänger – als „Hauptrisikogruppe im Straßenverkehr“ (S. 20) bezeichnen. Im Jahr 2009 waren 19 % aller Getöteten sowie jeder fünfte im Straßenverkehr Verletzte ein Mitglied dieser Altersgruppe, wobei ihr Anteil an der Gesamtbevölkerung in Deutschland lediglich 8,3 % betrug. Junge Fahranfänger sind damit trotz insgesamt rückläufiger Zahlen immer noch besonders gefährdet, bei einem Verkehrsunfall zu verunglücken. Der Anteil junger Fahranfänger, die im Straßenverkehr verunfallen, ist unverhältnismäßig hoch (z.B. Statistisches Bundesamt, 2006, 2008). In den ersten sechs Monaten, in denen die Führerscheinneulinge selbstständig fahren, ist die Unfallgefährdung für sie am höchsten (Mayhew, Simpson & Pak, 2003; Sagberg, 1998). Dieses erhöhte Unfallrisiko sinkt innerhalb von zwei Jahren nach dem Erwerb der Fahrerlaubnis deutlich (vgl. Abbildung 2.2); innerhalb der ersten neun Monate bereits um die Hälfte (Schade, 2001). Im weiteren Verlauf zeigt sich eine asymptotische Annäherung der Unfallrate an ein relativ geringes eher stabiles Niveau. Aufgrund des erhöhten Unfallrisikos der jungen Fahranfänger wurden in zahlreichen Studien Unfallursachen bei dieser Gruppe von Fahrern eingehend untersucht. Die Analyse von Unfallursachen und die Identifikation besonders häufig vorkommender Unfallarten gibt Aufschluss über mögliche Defizite beim Fahren. Daher werden im Folgenden Unfallursachen, insbesondere bei jüngeren Fahrern, dargestellt.

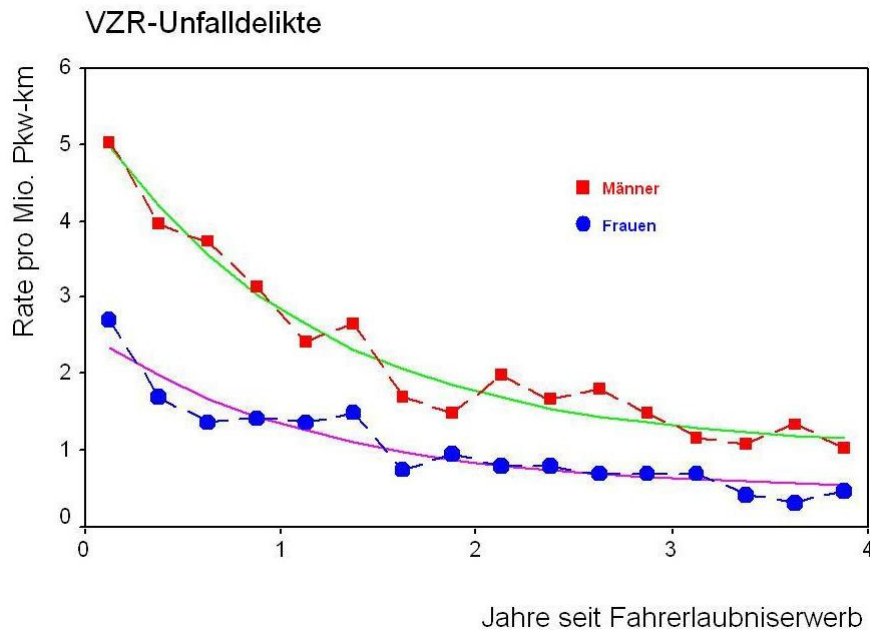


Abbildung 2.2

Unfälle mit Delikteintragung im Verkehrszentralregister von Fahranfängern in den ersten vier Jahren ihrer Fahrerkarriere (Jahr des Fahrerlaubniserwerbs: 1987; Männer: N=5205; Frauen: N=6095; aus Schade, 2001, S. 3)

Einer *In-Depth*-Unfallanalyse zufolge, die im Jahre 2002 in Braunschweig durchgeführt wurde, kommen insgesamt besonders viele Verkehrsunfälle durch nicht an die Situation angepassten Abstand zu anderen Verkehrsteilnehmern und zu hohe Geschwindigkeit zustande (Vollrath, 2010). Zudem konnte eine hohe Zahl an Unfällen dadurch erklärt werden, dass andere vorfahrtberechtigte Verkehrsteilnehmer nicht oder zu spät erkannt wurden.

In einer Studie von McKnight und McKnight (2003) wurden insgesamt 2000 Unfallberichte der Polizei begutachtet. In alle Unfälle waren zur einen Hälfte ganz junge Fahrer (16-17 Jahre) und zur anderen Hälfte Fahrer mit wenig Fahrerfahrung (18-19 Jahre) verwickelt. Die Fehler, die hauptsächlich als Ursache für die Unfälle betrachtet wurden, waren Fehler beim Beobachten des vorausfahrenden Verkehrs, des Verkehrs parallel zum eigenen Fahrzeug und des rückwärtigen Verkehrs (43 %), Unaufmerksamkeit (23 %) und unangepasste Geschwindigkeit (21 %). Dieses Muster unterschied sich nicht bedeutsam zwischen den beiden Altersgruppen. Klauer et al. (2006) konnten zeigen, dass die Rate von Unfällen, die durch Unaufmerksamkeit verursacht wurden, durchschnittlich etwa fünfmal höher bei Fahrern zwischen 18 und 20 Jahren ist als bei älteren Fahrern. Deutsche Unfallstatistiken zeigen zudem, dass unangemessene Geschwindigkeit, zu geringer Abstand und unangebrachtes Überholen zu den Hauptunfallursachen junger Fahranfänger gehören (Statistisches Bundesamt, 2008). Unangepasste Geschwindigkeit gehört insgesamt seit vielen

Jahren zu den hauptsächlichen Bedingungen für Unfälle im Straßenverkehr (z.B. Grattenthaler & Krüger, 2009).

2.2.1 Risikofaktoren

In Kapitel 2.1 wurde dargestellt, dass sich Experten in einer Domäne durch hohe Leistungen auszeichnen. Die unverhältnismäßig hohe Unfallrate junger Fahranfänger lässt auf eher schlechte Leistungen beim Autofahren und somit auf mangelnde Expertise schließen. Kurz nach dem Erwerb ihrer Fahrerlaubnis sind Fahranfänger besonders gefährdet. Die meisten Modelle des Expertiseerwerbs würden die Überlegenheit der erfahrenen Fahrer sowie den Abfall der Unfallrate in der ersten Zeit nach dem Führerscheinwerb mit einem Anstieg an Expertise durch Erfahrung erklären. Allerdings kann diese Erklärung nicht eindeutig getroffen werden, da sich erfahrene Fahrer von unerfahrenen Fahrern nicht nur durch die Menge an Fahrerfahrung unterscheiden: Unerfahrene Fahrer sind in der Regel im Durchschnitt deutlich jünger als erfahrene Fahrer.

Daher wird das erhöhte Unfallrisiko junger Fahrer auf zwei Ursachenkomplexe zurückgeführt die beide bei der Unfallentwicklung zusammen wirken: das Anfänger- und das Jugendlichkeitsrisiko (z.B. Leutner et al., 2009). Das heißt Fahranfänger bringen sich aufgrund ihrer Jugendlichkeit oft in Situationen im Straßenverkehr, die sie wegen mangelnder Erfahrung noch nicht meistern können. Beide Ursachenkomplexe können mit unterschiedlichen Defiziten der Fahranfänger in Verbindung gebracht werden.

2.2.1.1 Jugendlichkeitsrisiko

Junge Fahrer insgesamt haben ein höheres Verkehrsunfallrisiko als Fahrer mittleren Alters, wobei junge männliche Fahrer ein besonders hohes Risiko aufweisen (Evans, 1991; Laapotti, Keskinen, Hatakka & Katila, 2001; Schade, 2001; Williams, A. F., 2003). Dieses erhöhte Risiko wird mit unerwünschten Eigenschaften, Einstellungen und einem speziellen dysfunktionalen Lebensstil junger Fahrer in Verbindung gebracht (z.B. Elander, West & French, 1993; Keskinen, 1996; Laapotti et al., 2001; Williams, A. F., 2003) und oftmals als Jugendlichkeitsrisiko bezeichnet. Für das Jugendlichkeitsrisiko relevante Aspekte können durch die Erfassung des Fahrverhaltens, der Unfallarten und des Fahrerlebens junger Fahrer identifiziert werden.

Laapotti et al. (2006) führten in Verbindung mit dem EU-Projekt Basic (Hatakka et al., 2003) eine repräsentative Studie mit 7789 Teilnehmern durch, in der Selbstauskünfte von Fahranfängern und Fahrern im mittleren Alter über ihr Fahrverhalten und ihre

Unfallgeschichte erfasst wurden. Fahranfänger berichteten darin z.B. häufiger bei Nacht zu fahren als Fahrer im mittleren Alter. Kroj und Schulze (2002) berichten, dass besonders viele tödliche Unfälle junger Fahranfänger bei nächtlichem Freizeitverhalten, z.B. nach Discobesuchen erfolgen.

Cooper, Pinili und Chen (1995) untersuchten die Unfälle von Fahranfängern zwischen 16 und 55 Jahren. Da sämtliche Fahrer über eine ähnlich hohe Fahrerfahrung verfügten, konnten Unterschiede bei der Unfallbeteiligung eindeutig auf das Alter zurückgeführt werden. Es zeigte sich, dass die Unfälle jüngerer Fahrer häufiger mit Beifahrern zusammen stattfanden.

Unangepasste Geschwindigkeit gehört zu den Hauptunfallursachen im Straßenverkehr (z.B. Grattenthaler & Krüger, 2009). Rasen (*Speeding*) wurde konsistent als Prädiktor für Unfallanfälligkeit ermittelt (z.B. Horswill & McKenna, 1999; Quimby, Maycock, Carter, Dixon & Wall, 1986). Trotz des hohen Risikos, das mit überhöhter Geschwindigkeit einhergeht, waren Geschwindigkeitsüberschreitungen auch im Jahre 2009 wieder häufigster registrierter Verkehrsverstoß für Männer und Frauen (Kraftfahrt-Bundesamt, 2010). Unfallstatistiken zeigen, dass besonders viele Unfälle bei Fahranfängern aufgrund unangepasster Geschwindigkeit zustande kommen (z.B. McKnight & McKnight, 2003). Die Wahl einer unangepassten Geschwindigkeit kann dem Jugendlichkeitsrisiko zugeordnet werden, da dieser Aspekt nicht von kognitiven oder psychomotorischen Fähigkeiten abhängen sollte, die ein Fahrer durch Erfahrung erwirbt. Viel eher sollten bei der Entscheidung zur Geschwindigkeitsübertretung motivationale Aspekte überwiegen. Besonders diejenigen Fahranfänger, die bereits etwas Fahrerfahrung gesammelt haben, tendieren dazu, deutlich zu schnell zu fahren (Ellinghaus & Steinbrecher, 1990). Dieser Befund könnte mit den Erkenntnissen von McKenna und Horswill (2006) und auch Moe (1996) begründet werden, die zeigen konnten, dass die Wahl der Fahrtgeschwindigkeit positiv mit der Einschätzung der eigenen Fahrfähigkeiten zusammenhängt. Svenson (2009) zeigte zudem, dass die Steigerung des Risikos eines schwerwiegenden Unfalls, die mit einer Erhöhung der Fahrtgeschwindigkeit einhergeht, unterschätzt wird. Es zeigte sich außerdem, dass die Geschwindigkeitswahl bei jungen unerfahrenen Fahrern besonders konstant über verschiedene Verkehrssituationen hinweg ist (Grayson, Maycock, Groeger, Hammond & Field, 2003). Dieser Befund weist darauf hin, dass junge Fahrer nicht zwischen Situationen mit unterschiedlichem Gefahrenpotenzial differenzieren. Sie passen ihre Geschwindigkeit nicht an die Anforderungen der aktuellen Situation an.

Das Vertrauen auf die eigenen Fahrfertigkeiten ist demnach ein Aspekt, der mit dem Fahrverhalten interagiert. Eine überhöhte Einschätzung eigener Fahrfähigkeiten wird häufig in der Literatur als ein Faktor betrachtet, der mit einem erhöhten Unfallrisiko junger Fahrer in Zusammenhang gebracht wird (z.B. Gregersen, 1996; Keskinen, 1996). Studien, die Zusammenhänge zwischen der absoluten Unfallbeteiligung und der Selbsteinschätzung untersuchen, erbrachten jedoch widersprüchliche Ergebnisse, wobei sich nur in einigen Studien ein positiver Zusammenhang zwischen Selbsteinschätzung und Unfallverwicklung zeigen lässt (vgl. Biermann, 2007). Es konnte gezeigt werden, dass ein großer Anteil finnischer (ca. 30 %), schwedischer (58 - 70 %) und niederländischer (50 %) Fahranfänger ihre Fahrkompetenzen (z.B. Fahrzeugsteuerung, ökonomisches Fahren, Verkehrssicherheit, Gefahrenerkennung, Verkehrsfluss, Verhalten gegenüber schwächeren Verkehrsteilnehmern und Kontrolle über Verkehrssituationen) im Vergleich zu dem Urteil eines Fahrprüfers überschätzten (Mynttinen, 2010). Allerdings ist anzumerken, dass auch ein großer Anteil der Fahranfänger in der Lage war, die eigenen Fähigkeiten adäquat einzuschätzen. Es scheint demnach bei Fahranfängern nicht generell zu einer Falscheinschätzung der eigenen Fahrkompetenz zu kommen.

Ein relativ gesicherter Befund scheint zu sein, dass Fahrer ihre eigenen Fähigkeiten mit steigender Fahrerfahrung höher einschätzen (Groeger, 2000; Spolander, 1982). Eine noch offene Frage hierbei ist, ob sich die Entwicklung der Selbsteinschätzung auch im Längsschnitt zeigen kann. Die Angst vor einem Unfall nimmt in den ersten Monaten nach dem Fahrerlaubniserwerb ab, was auf ein zunehmendes Vertrauen in die eigenen Fähigkeiten hindeutet (Ellinghaus & Steinbrecher, 1990). Die anfänglichen Ängste sollten sich, wie spezifische Ängste im Allgemeinen (Saß, Wittchen, Zaudig & Houben, 2003), in negativen Emotionen gegenüber dem Fahren, Sorgen sowie in Vermeidungsverhalten äußern. Problematisch dabei ist, dass Ängste und daraus resultierendes Vermeidungsverhalten den Aufbau von Fahrkompetenz und den Abbau von Ängsten hinsichtlich des Fahrens verlangsamen oder gar verhindern könnten. Zur Entwicklung des Angstempfindens zu Beginn der Fahrerkarriere besteht jedoch noch Forschungsbedarf.

2.2.1.2 Anfängerrisiko

Zu Beginn der Fahrkarriere ist unabhängig vom Lebensalter des Fahranfängers ein erhöhtes Unfallrisiko zu verzeichnen. Belege für dieses generelle Anfängerrisiko erbrachten z.B. Maycock, Lockwood und Lester (1991), die Fahranfänger in verschiedenen Altersgruppen untersuchten. Die Autoren trennten, basierend auf einer Studie mit einer Stichprobe von

18 500 Fahrern in Großbritannien, Anfänger- und Jugendlichkeitsrisiko glaubwürdig voneinander. Diese Trennung war möglich, da das Alter der Personen und die Dauer des Fahrerlaubnisbesitzes innerhalb der Stichprobe systematisch variierten. In der Studie wurde erfasst, wie oft die Befragten in den letzten drei Jahren (bzw. seit Erhalt der Fahrerlaubnis) in Unfälle verwickelt waren und um welche Art von Unfällen es sich dabei handelte. Basierend auf den erhobenen Daten konnte berechnet werden, wie hoch die zu erwartende Unfallhäufigkeit innerhalb eines Jahres (bei einer Fahrleistung von 7 500 Meilen) ist. Die jährliche Unfallhäufigkeit variierte dabei in Abhängigkeit von Alter und Fahrerlaubnisbesitzdauer. In Abbildung 2.3 wird durch die Darstellung der Reduktion des anfänglichen Unfallrisikos mit zunehmendem Alter der Fahreinsteiger das Jugendlichkeitsrisiko (strichlinierte Kurve) verdeutlicht. Je älter die Fahranfänger beim Erwerb ihrer Fahrerlaubnis waren, desto geringer war ihre zu erwartende Unfallhäufigkeit unmittelbar nach dem Erhalt der Fahrerlaubnis. Ungeachtet dessen, mit welchem Alter die Fahrer ihre Fahrerlaubnis erwerben, reduziert sich die erwartete Unfallhäufigkeit mit der Zeit, was auf den durch Erfahrung geringer werdenden Einfluss des Anfängerrisikos hindeutet (durchgezogene Kurven). Das Unfallrisiko verringert sich negativ beschleunigt in jeder Altersgruppe, aber auch über die Altersgruppen hinweg. Das bedeutet, dass zu Anfang der Fahrkarriere das Unfallrisiko sehr schnell sinkt und sich mit zunehmender Erfahrung langsamer an einen Minimalwert annähert. Über die Gruppen hinweg zeigt sich, dass sich ein etwas höheres Einstiegsalter bei jüngeren Fahranfängern deutlicher positiver auf die Unfallhäufigkeit auswirken sollte, als bei älteren Fahranfängern. Zu bedenken ist allerdings, dass es sich bei der Stichprobe älterer Fahranfänger um eine selektive Stichprobe handelt. Ob sich die Personen, die in die Studie einbezogen wurden, in relevanten Merkmalen von der Grundgesamtheit der Personen in dem entsprechenden Alter unterscheiden, wurde in dieser Studie nicht erfasst. Zu ähnlichen Ergebnissen gelangten auch Mayhew, Simpson und Pak (2003) in einer kanadischen Studie, bei der sie Unfallberichte aus den Jahren 1992 und 1993 analysierten. Sie untersuchten Veränderungen in den Unfallraten junger Fahrer von Monat zu Monat über die ersten beiden Jahre ihres Führerscheinbesitzes hinweg. Dabei zeigte sich, dass die Unfallzahlen innerhalb der ersten sechs Monate deutlich fielen. Dieses Ergebnis lässt auf einen deutlichen Einfluss des Anfängerrisikos schließen, da sich Einflüsse, die im Zusammenhang mit dem Alter einer Person stehen, innerhalb von sechs Monaten kaum ändern sollten. Im Gegensatz dazu sammelten die Fahranfänger mit jeder Fahrt mehr Erfahrung, was ihr Anfängerrisiko reduziert haben sollte.

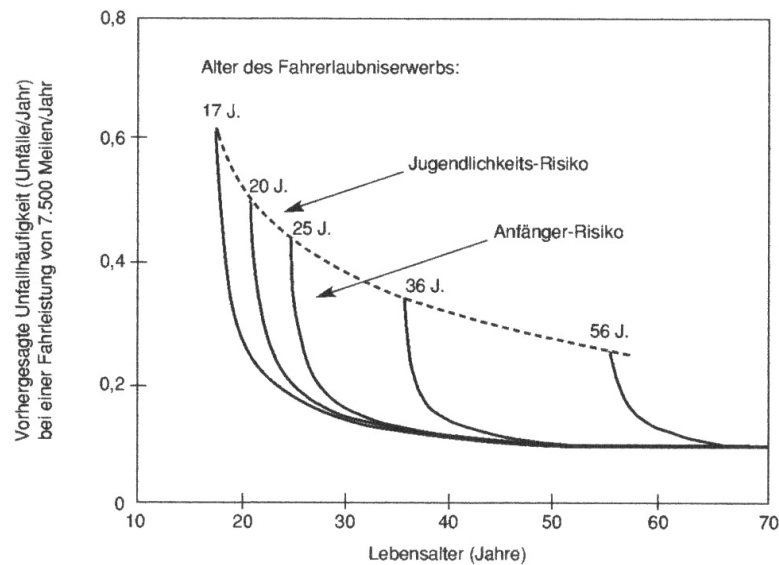


Abbildung 2.3

Darstellung von Anfänger- und Jugendlichkeitsrisiko anhand der vorhergesagten jährlichen Unfallraten für Fahranfänger mit unterschiedlichem Einstiegsalter (nach Maycock, Lockwood, & Lester, 1991; aus Leutner, Brünken, & Willmes-Lenz, 2009, S.2)

Ballesteros und Dischinger (2002) fanden in ihrer Analyse der Unfalldaten, die zwischen 1996 und 1998 in Polizeiberichten registriert wurden, ebenfalls Hinweise auf das Anfängerrisiko. Dabei wurden Unfallanalysen für die Fahrer von 16 bis 21 Jahren getrennt in Schritten zu je einem Jahr vorgenommen. Das Alter der Personen beim Führerscheinerwerb wurde in dieser Studie nicht einbezogen. Es zeigte sich, dass die 16-Jährigen im Verhältnis zur gefahrenen Kilometerzahl die meisten Unfälle verursacht hatten. Die Unfallzahlen sanken mit steigendem Alter. Die Autoren nahmen an, dass das hohe Risiko der jüngsten Fahrer eher auf mangelnde Erfahrung zurückzuführen sei. Diese Annahme resultierte aus der Beobachtung, dass die 16-Jährigen im Vergleich zu den anderen Altersgruppen mehr Unfälle bei „sicheren“ Bedingungen verursachten (z.B. bei Tag und gutem Wetter). Die Älteren wiesen hingegen im Verhältnis mehr Unfälle auf, die in Verbindung mit risikoreichen Verhaltensweisen (z.B. Alkohol- und Drogenkonsum) standen. Dabei war das Risiko der jungen Männer und Frauen mit 16 Jahren gleich hoch, während bei den älteren Gruppen die männlichen Fahrer mehr Unfälle verursachten.

McKnight und McKnight (2003) arbeiteten bei ihrer Studie ebenfalls mit Unfallberichten junger Fahrer in Maryland und Kalifornien. Dabei konnten sie zeigen, dass proportional mehr Unfälle der jüngsten Fahrer auf das Anfängerrisiko zurückzuführen waren. Sie begründeten dies analog zu Ballestero und Dischinger (2002) damit, dass die jüngeren

Fahrer mehr Unfälle aufgrund mangelnder Basisfertigkeiten beim Fahren verursachten und die älteren mehr aufgrund von Risikoverhaltensweisen.

2.2.1.3 Anteile von Jugendlichkeits- und Anfängerrisiko am Gesamtrisiko

In den Kapiteln 2.2.1.1 und 2.2.1.2 wurden zwei Ursachenkomplexe vorgestellt, die das erhöhte Unfallrisiko junger Fahrer erklären sollen: das Jugendlichkeits- und das Anfängerrisiko. Besonders im Hinblick auf Interventionsmaßnahmen, wenn eine Entscheidung darüber ansteht, welche Aspekte des Fahrverhaltens junger Fahrer man besser trainieren sollte, stellt sich die Frage, welcher der beiden Risikoaspekte mehr Einfluss ausübt. Es gibt sowohl Autoren, die das Anfängerrisiko als entscheidender erachten (z.B. Leutner et al., 2009; Stiensmeier-Pelster, 2007) als auch andere, die dem Jugendlichkeitsrisiko gleichviel (Mayhew & Simpson, 1996) oder mehr Bedeutung (Jonah, 1986) einräumen. Unfallanalysen haben gezeigt, dass das Unfallrisiko in den ersten Monaten nach dem Fahrerlaubniserwerb deutlich sinkt (Maycock et al., 1991; Mayhew et al., 2003; Schade, 2001). Dieser Befund deutet auf einen höheren Einfluss des Anfängerrisikos hin, da innerhalb dieses kurzen Zeitraums, ein Anstieg an Erfahrung, nicht aber eine deutlichere Verringerung der Jugendlichkeit der Fahranfänger zu erwarten ist. Allerdings scheint auch das Lebensalter einen Einfluss auf das Unfallrisiko zu haben. Maycock, Lockwood und Lester (1991) konnten zeigen, dass Fahranfänger weniger Unfälle zu Beginn ihrer Fahrkarriere machen, je älter sie beim Erwerb ihrer Fahrerlaubnis sind. Ein deutlicher Abfall der Unfallrate innerhalb der ersten Monate zeigte sich jedoch für alle Altersklassen.

Denkbar wäre, dass sich die Anteile von Jugendlichkeits- und Anfängerrisiko am Gesamtunfallrisiko einer Person mit der Zeit ändern. Genauere Analysen der Unfälle junger Fahrer haben gezeigt, dass sich die Unfälle, die ganz zu Beginn des eigenständigen Fahrens verursacht werden, in ihrer Art von den Unfällen unterscheiden, die später passieren, wenn die Personen zwar immer noch jung aber schon etwas erfahrener sind. In den ersten Monaten handelt es sich eher um Unfälle, die auf mangelnde Erfahrung zurückzuführen sind, während der Anteil an Unfällen, die aus risikoreichem Fahren resultieren, später ansteigt (Ballesteros & Dischinger, 2002; McKnight & McKnight, 2003).

Man kann demnach schlussfolgern, dass das Anfängerrisiko in den ersten Monaten nach dem Fahrerlaubniserwerb das Jugendlichkeitsrisiko deutlich übertrifft. Danach sollte der Einfluss des Lebensalters einen höheren Anteil am Gesamtrisiko haben, als kurz nach dem Fahrerlaubniserwerb.

2.2.2 Entwicklung von Fahrkompetenz

Aus Sicht der Pädagogischen Psychologie kann der Aufbau von Fahrkompetenz als ein Prozess verstanden werden, bei dem verschiedene kognitive, affektive und psychomotorischer Teilfertigkeiten erworben werden (Leutner et al., 2009). Zur Entwicklung von Fahrkompetenz oder Fahrexpertise liegen bisher zwar Modelle vor, die jedoch noch nicht ausreichend empirisch abgesichert sind. Innerhalb der Fahrausbildung gestaltet sich der Beginn der Expertiseentwicklung für alle Fahranfänger sehr ähnlich. Durch den Aufbau der formalisierten Fahrausbildung in Deutschland bedingt, erfolgt zunächst innerhalb des theoretischen Fahrschulunterrichtes die Aneignung von verkehrsspezifischem Faktenwissen bezüglich Verkehrsregeln, Verkehrsverhalten, Gefahren im Verkehr, rechtlichen Rahmenbedingungen, Fahrzeugtechnik und Fahrphysik, persönlichen Voraussetzungen und dem Verhalten anderer Verkehrsteilnehmer (§1 Fahrschüler-Ausbildungsordnung [FahrschAusbO])². Im praktischen Fahrschulunterricht, der entweder nach der theoretischen Fahrerlaubnisprüfung oder parallel zum theoretischen Unterricht absolviert wird, lernen die Fahrschüler die praktische Fahrzeughandhabung und erproben das im Theorieunterricht erlernte Faktenwissen im Realverkehr. Am Ende der Fahrausbildung haben die Fahrschüler demnach verkehrsspezifisches Faktenwissen, die Fahrzeugbeherrschung (in Standardsituationen) sowie grundlegende Fertigkeiten zur Fahrzeughandhabung erworben.

Das hohe Einsteigerrisiko von Fahranfängern zeigt allerdings, dass das erworbene Wissen und die Fähigkeiten scheinbar nicht ausreichen, um sich sicher im Verkehr bewegen zu können. Wie in Kapitel 2.2.1 dargestellt, zeigen Unfallstatistiken, dass die Fahranfänger durch Erfahrung beim eigenständigen Fahren zunehmend Expertise aufbauen, wodurch sich ihr Unfallrisiko verringert. Wie sich andere Teilaspekte des Fahrenkönnens während der ersten Zeit des eigenständigen Fahrens entwickeln, ist bisher jedoch nicht bekannt. Allerdings kann vermutet werden, dass sich zunächst Aspekte des Anfängerrisikos verändern, da sich die Defizite in diesem Bereich durch Erfahrung reduzieren lassen.

Zur Identifikation derjenigen Fähigkeiten und Fertigkeiten, die in der ersten Zeit nach dem Fahrerlaubniswerb erworben werden, kann ein hierarchisches Modell des Fahrverhaltens der Arbeitsgruppe um Keskinen (Hatakka, Keskinen, Gregersen & Glad, 1999; Keskinen, 1996) dienen. Laut diesem Modell kann das Fahrverhalten in vier Ebenen zunehmender Komplexität aufgespalten werden (vgl. Abbildung 2.4).

² Fahrschüler-Ausbildungsordnung vom 18. August 1998 (BGBl. I S. 2307, 2335), zuletzt geändert durch Artikel 3 der Verordnung vom 18. Juli 2008 (BGBl. I S. 1338).

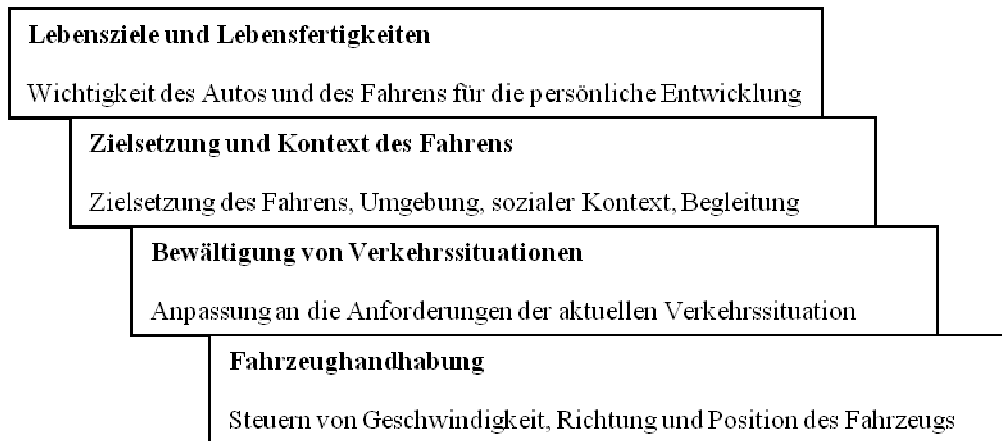


Abbildung 2.4
Hierarchisches Modell des Fahrverhaltens (nach Hatakka et al., 1999)

Die beiden höchsten Ebenen beinhalten motivationale, affektive und emotionale Aspekte beim Fahrer, die auf das Fahrverhalten einwirken. Auf der obersten Ebene „Lebensziele und Lebensfertigkeiten“ spielen selbstregulative Prozesse und Lebensziele (allgemeine Fähigkeiten, Persönlichkeitseigenschaften und Einstellungen) eine Rolle, die sich nicht nur auf die aktuelle Fahrt beziehen, sondern eher genereller Natur sind. Auf der nächst unteren Ebene „Zielsetzung und Kontext des Fahrens“ wird der Kontext der aktuellen Fahrt wichtig. Dies schließt einerseits Stimmung (z.B. Stress, Euphorie) und Motive (z.B. zum Spaß fahren, zur Schule fahren) des Fahrers, sowie die äußeren Umstände der Fahrt ein (z.B. Beifahrer). Die beiden oberen Ebenen beinhalten Aspekte, die im Zusammenhang mit dem Jugendlichkeitsrisiko als kritisch betrachtet werden.

Auf den beiden unteren Ebenen finden sich Aspekte der Aneignung von Fahrhandlungsroutine. Die Ebene „Bewältigung von Verkehrssituationen“ bezeichnet Fähigkeiten beim Agieren im Straßenverkehr, die gewährleisten, dass sich der Fahrer in einer Verkehrssituation angemessen verhält. Als bedeutsam werden auf dieser Ebene unter anderem auch das Wissen um Verkehrsregeln sowie die Fähigkeit, Verkehrssituationen angemessen einschätzen zu können, erachtet. Auf der Ebene „Fahrzeughandhabung“ werden Fertigkeiten eingeschlossen, die die Kontrolle über das Fahrzeug beeinflussen. Darunter sind zum Beispiel die Spurhaltung und die Geschwindigkeitsregulation zu verstehen. Die dargestellten Aspekte der beiden unteren Ebenen wirken sich auf das Anfängerrisiko aus. Es kann gefolgert werden, dass sich durch Erfahrung Fahroutine aufbaut. Wie routiniertes Fahrverhalten entsteht, kann auf der Grundlage des Informationsverarbeitungsmodells von Atkinson und Shiffrin (1969) erklärt werden (Leutner & Brünken, 2002). Generell übertreffen Experten Novizen auf einem spezifischen Fachgebiet, auf dem sie durch Übung viel Wissen und Können ausgebildet haben

(vgl. 2.1.2). In einer spezifischen Domäne machen Experten weniger Fehler und lösen spezifische Probleme schneller als Novizen (Gruber, 1994). Die Überlegenheit von geübten Fahrern im Gegensatz zu Fahranfängern bei Routineaufgaben im Straßenverkehr kann dadurch erklärt werden, dass Teilhandlungen des Fahrens bereits automatisiert sind. Rasmussen (1986) hat auf der Grundlage dieser Annahme ein Modell der Kontrolle und Regulation von Handlungen auf drei Ebenen entwickelt (vgl. Abbildung 2.5). Die routinemäßigen Handlungen beim Fahren nehmen kaum Ressourcen des Arbeitsgedächtnisses in Anspruch. So kann der routinierte Fahrer seine Aufmerksamkeit auf diejenigen Aspekte des Fahrens richten, die dieser noch bedürfen.

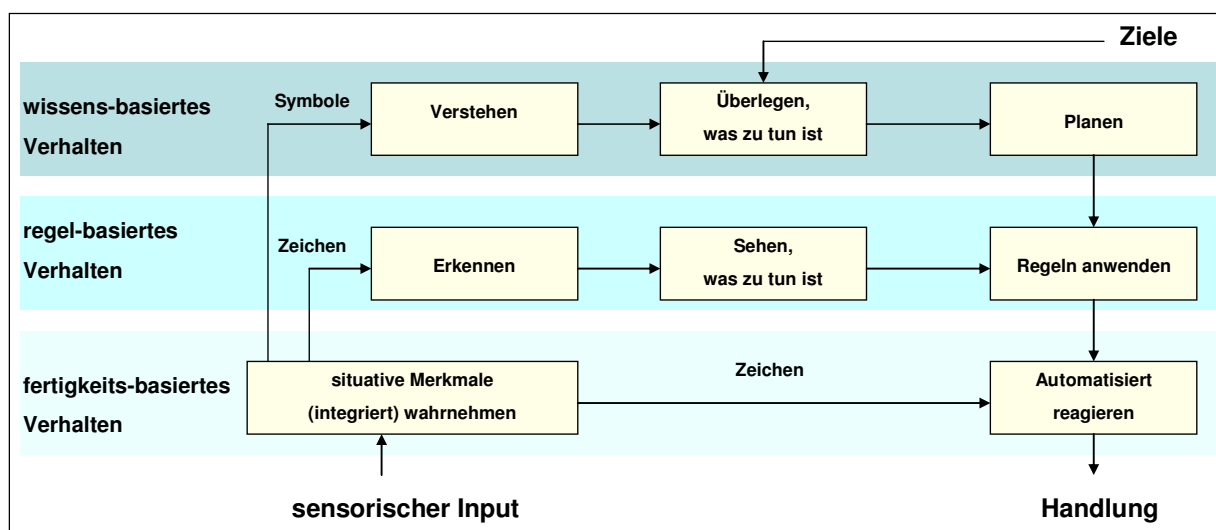


Abbildung 2.5

Vereinfachtes Modell der Handlungskontrolle auf drei Ebenen (übersetzt von Leutner & Brünken, 2002; nach Rasmussen, 1986)

Leutner und Brünken (2002) beziehen dieses Modell auf das zielgerichtete Handeln beim Autofahren mit seinen verschiedenen interagierenden Teilhandlungen (vgl. auch Bartmann, Debus & Heller, 1994; Gregersen et al., 2000).

Gemäß Rasmussens Modell erfolgen Teilhandlungen ihrem Automatisierungsgrad entsprechend auf einer von drei Ebenen der Handlungskontrolle: der wissensbasierten Ebene, der regelbasierten und der fertigkeitbasierten Ebene. Rasmussen (1986) weist darauf hin, dass die drei Ebenen in der Realität auf eine komplexere Weise miteinander interagieren als im Modell dargestellt. Die Pfeile repräsentieren im in Abbildung 2.5 dargestellten Modell der Handlungskontrolle den Informationsfluss.

Die wissensbasierte Ebene stellt diejenige Ebene dar, auf der Handlungen in unvertrauten Situationen reguliert werden. Das Ziel der Handlung ist dabei explizit formuliert

und steuert die Handlung direkt. Fahrschüler lernen die grundlegenden Fertigkeiten des Fahrzeughandlings relativ schnell (Hall & West, 1996), automatisiert erfolgt die Fahrzeughandhabung dennoch am Anfang noch nicht. Der Weg vom sensorischen Input bis zur entsprechenden Handlung ist noch sehr lang. Eingehende Reize, also die Merkmale der aktuellen Verkehrssituation, müssen zunächst interpretiert und verstanden werden. Anschließend wird bewusst überlegt, was als nächstes zu tun ist, um dem Ziel näher zu kommen. Ganz bewusst erfolgt auch die Handlungsplanung, wobei eine passende Regel zunächst ausgewählt und adäquat angewandt werden muss, bevor die Handlungsausführung letztendlich erfolgt. Ein Fahrer, der sich auf der wissensbasierten Ebene befindet, wird vermutlich den Drehzahlmesser im Armaturenbrett beobachten. Er wird sich an die Regel erinnern, dass ab einer bestimmten Drehzahl in den nächsten Gang geschaltet werden muss. Demnach wird er sich vornehmen, etwas Gas zu geben bis die Nadel den kritischen Bereich erreicht hat und daraufhin den Vorgang des Schaltens einzuleiten. Die einzelnen Schritte wird er ganz bewusst durchführen.

Auf der regelbasierten Ebene wird die Handlung nicht mehr direkt durch das Ziel gesteuert, sondern durch im Gedächtnis gespeicherte Regeln. Der Fahrer erkennt die Reizgegebenheit sofort als „Zeichen“, welche Handlungen erforderlich sind. Er handelt nach den Regeln, die er bewusst aus seinem Handlungsrepertoire auswählt. Auf dieser Stufe muss der Fahrer den Drehzahlmesser nicht beobachten. Am Geräusch des Motors erkennt er bereits, dass das Schalten in den nächsten Gang erforderlich ist. Das Schalten erfolgt noch immer bewusst.

Auf der Stufe, bei der der Fertigkeitserwerb abgeschlossen ist, der fertigkeitbasierten Ebene, wird nach dem sensorischen Input von außen sofort eine zielgerichtete Handlung eingeleitet. Diese Handlung läuft demnach automatisiert ab. Automatisierte Routinen basieren dabei auf der unbewussten, zeitlichen und räumlichen Verarbeitung von Objekten und Symbolen in einem bekannten Umfeld. Dem geübten Fahrer ist die Entscheidung über das Schalten, sowie das Schalten selbst nicht bewusst. Die einzelnen Schritte, die dabei kognitiv und motorisch abgelaufen sind, können kaum noch erinnert werden. Personen, die sich hinsichtlich der Fahrzeugführung auf der höchsten Fertigkeitsebene befinden, müssen beim Fahren nicht mehr bewusst entscheiden, wann sie in den nächsten Gang schalten. Diese Teilhandlung erfolgt unbewusst und somit automatisiert. Aus diesem Modell der Handlungskontrolle kann abgeleitet werden, dass Personen, deren Entscheidungen und Handlungen beim Fahren noch nicht automatisch ausgeführt werden, diese in bewussten

Teilschritten aufgegliedert durchführen müssen. Dabei müssen auch jeweils relevante Regeln beim Fahren bewusst erinnert und umgesetzt werden.

Rasmussens Modell kann analog zu dem dreistufigen Modell von Fitts und Posner (1967) oder dem in Kapitel 2.1.3.1 vorgestellten ACT*-Modell von Anderson (1982, 1993) als Modell zum Erwerb von Expertise interpretiert werden. Fitts und Posner unterscheiden drei Phasen des Fertigkeitserwerbs: die kognitive Phase (deklarative Phase bei Anderson; wissensbasierte Ebene bei Rasmussen), die assoziative Phase (Phase der Wissenskompilierung bei Anderson; regelbasierte Ebene bei Rasmussen) und die autonome Phase (prozedurale Phase bei Anderson; fertigkeit-basierte Ebene bei Rasmussen), die beim Aufbau automatisierter Fertigkeiten durchlaufen werden können. Durch Training von Routineaufgaben erreicht man dabei die letzte Phase. Demnach kann das Modell von Rasmussen zum einen dem Verständnis des Zusammenspiels von Handlungen beim Fahren dienen, die (noch) auf unterschiedlichen Ebenen verarbeitet werden. Zum anderen beschreibt es aber auch den Aufbau von Automatisierung bei Routineaufgaben.

Wie in diesem Kapitel bereits erwähnt, beanspruchen bewusst ausgeführte Handlungen die knappen Ressourcen des Arbeitsgedächtnisses. Automatisierte Aufgaben hingegen sollten das Arbeitsgedächtnis nicht belasten (z.B. Leutner & Brünken, 2002). Automatisierung bildet sich erst mit zunehmender Fahrerfahrung aus. Daraus kann man schließen, dass das Fahren Fahranfänger kognitiv stärker beansprucht als Fahrexperten. Unterstützung für diese These liefern Studien, die das Doppeltätigkeitsparadigma einsetzen (Bartmann, 1995; Bartmann et al., 1994; Brünken, Plass & Leutner, 2003). Die Anfänger zeigten dabei etwas stärkere Leistungseinbußen bei Nebenaufgaben, was als eine gegenüber den Experten erhöhte kognitive Beanspruchung der Novizen interpretiert werden kann.

2.2.3 Zusammenfassung und Forschungsdesiderate

In der vorliegenden Arbeit wurde das Ziel verfolgt, ein geeignetes Modell zur Beschreibung des beginnenden Erwerbs der Fahrexpertise zu identifizieren. Dazu sollten Novizen beim Expertiseerwerb innerhalb einer Längsschnittstudie beobachtet werden. Nachdem in Kapitel 2.1 generelle Merkmale der Expertise, verschiedene Modelle des Expertiseerwerbs sowie die Besonderheiten der Fahrexpertise gegenüber der Expertise in anderen Domänen dargestellt wurden, fehlten zur Durchführung einer Längsschnittstudie noch relevante Teilfertigkeiten der Fahrexpertise, die dabei erhoben werden sollten. In Kapitel 2.2 wurden daher zur Identifikation relevanter Komponenten der Fahrexpertise, Unfallursachen von Fahrnovizen identifiziert, die Hinweise auf deren Defizite geben. Daraus lassen sich im Umkehrschluss

diejenigen Fähigkeiten ableiten, die ein Fahrexperte, der sich sicher im Verkehr bewegt, haben sollte. Diese Fertigkeiten sollten demnach beim Expertiseerwerb aufgebaut werden und demzufolge zur Abbildung des Erwerbsprozesses erhoben werden.

Bei der Analyse von Unfällen junger Fahranfänger hat sich gezeigt, dass besonders in der ersten Phase des eigenständigen Fahrens Risikofaktoren eine Rolle spielen, die im Zusammenhang mit der unzureichenden Fahrerfahrung (Anfängerrisiko) stehen. Besonders häufig kommt es, wegen mangelnder Anpassung der Geschwindigkeit und des Abstands zu vorausfahrenden Fahrzeugen an die aktuellen Verkehrsgegebenheiten, zu Unfällen. Problematisch scheint auch das Überholen zu sein. Das Anfängerrisiko wirkt sich besonders auf zwei Ebenen des Fahrverhaltens aus – die Fahrzeughandhabung und die Bewältigung von Verkehrssituationen – welchen daher zur Ermittlung relevanter Expertisemerkmale besondere Aufmerksamkeit geschenkt wird. Zur Untersuchung des Erwerbs von Fahrexpertise erscheint demnach relevant, wie sich Fertigkeiten zur Bewältigung der Anforderungen auf diesen Ebenen ausbilden. Dabei hat sich gezeigt, dass Rasmussens Modell den Aufbau von Expertise in Routinehandlungen erklärt und somit prinzipiell auf den Erwerb von Fahrexpertise bezogen werden kann. Sollen auf der Grundlage dieses Modells Vorhersagen bezüglich des Aufbaus von Fahrexpertise getroffen werden, ergeben sich dennoch Schwierigkeiten. Es ist noch nicht ausreichend empirisch überprüft, bei welchen Teilaspekten der Fahrkompetenz es sich tatsächlich um Routinehandlungen handelt und welche Teilhandlungen unabhängig vom Expertisegrad einer Person immer auf einer bewussten Ebene reguliert werden. Dass psychomotorische Prozesse beim Fahren in Standardsituationen automatisiert ablaufen ist unbestritten. Ob es bei Wahrnehmungsprozessen und in Ausnahmesituationen zu automatisiertem Verhalten kommt, ist jedoch ungewiss und eher unwahrscheinlich, da es sich dabei um ein seltenes und daher wenig geübtes Verhalten handelt, das infolge dessen nicht automatisiert vorliegen und auf einer unbewussten Ebene ausgeführt werden kann. Für den Erwerb von Expertise bezüglich dieser Teilhandlungen des Fahrens könnten andere Modelle des Expertiseerwerbs angemessener sein.

Die Lösung dieses Problems kann durch angemessene empirische Untersuchungsmethoden erfolgen. Studien, die Fahranfänger am Anfang ihrer Fahrkarriere über einen längeren Zeitraum begleiten, sind bisher höchst selten. Das Modell von Rasmussen (1986) wurde bisher nicht durch derartige Untersuchungen auf seine Gültigkeit für den Erwerb von Fahrexpertise validiert. Es scheint aber für Aspekte der Ebene Fahrzeughandhabung, auf der Routinehandlungen eine Rolle spielen, angemessen zu sein.

Allerdings werden Fertigkeiten auf dieser Ebene relativ schnell erworben (Hall & West, 1996), weshalb sie nicht als besonders kritisch für das Unfallrisiko betrachtet werden können. Wesentlicher scheinen Fertigkeiten zu sein, die die Bewältigung von Verkehrssituationen ermöglichen. Wie sich Aspekte dieser Ebene entwickeln, ist bisher nicht ausreichend empirisch geklärt. Um dies innerhalb dieser Arbeit zu untersuchen, muss zunächst aufgearbeitet werden, welche Aspekte dieser Ebene sicherheitsrelevant sind und wie sie operationalisiert werden können.

Im Folgenden wird ein sicherheitsrelevanter Faktor bei der Bewältigung von Verkehrssituationen – die Fähigkeit zur Gefahrenwahrnehmung – vorgestellt, bei der automatisierte und nicht automatisierte Teilhandlungen zur Gewährleistung einer sicheren Fahrweise interagieren.

2.3 Gefahrenwahrnehmung

Nach Horswill und McKenna (2004) handelt es sich bei der Schlüsselfrage der Wissenschaftler, die sich mit Verkehrssicherheit beschäftigen, darum, ob es quantifizierbare Komponenten der Fahrexpertise gibt, die erklären können, weshalb einige Fahrer häufiger in Unfälle verwickelt sind als andere Fahrer. Die Identifikation dieser Komponenten könnte sich positiv auf die Fahrsicherheit junger Fahrer auswirken, da man die Fahrausbildung und die Fahrerlaubnisprüfung entsprechend anpassen könnte. Dadurch würden risikorelevante Fähigkeiten besser trainiert und gemessen werden können. Im Folgenden wird ein Konzept vorgestellt – die Gefahrenwahrnehmung – das die erläuterten Anforderungen zu erfüllen scheint.

Es hat sich gezeigt, dass viele Unfälle junger Fahrer auf unzureichendes Absuchen der Umwelt auf potenzielle Gefahrenreize und Unaufmerksamkeit zurückzuführen sind (Klauer, S. G. et al., 2006; McKnight & McKnight, 2003). Erfahrene Fahrer scheinen unerfahrene Fahrer in der Fähigkeit zu übertreffen, den Verkehr aufmerksam nach Gefahren abzusuchen und komplexe Verkehrssituationen schnell zu überblicken. Diese Fähigkeit wird auch als Gefahrenwahrnehmung (engl. *Hazard Perception*) bezeichnet. McKenna und Crick (1994b, S. 1) schlagen für das Konzept der Gefahrenwahrnehmung folgende Definition vor: „Hazard perception refers to the ability to identify potentially dangerous traffic situations“. Als Gefahrenwahrnehmung kann man demnach die Fähigkeit bezeichnen, Situationen im Verkehr, die sich zu Gefahrensituationen entwickeln können, zu erkennen. Mills, Hall, McDonald und Rolls (1998) nennen dies auch die Fähigkeit, die Straße zu *lesen*. Dies

beschreibt anschaulich den Vorgang, aus den zahlreichen Sinneseindrücken einer komplexen Verkehrssituation sicherheitsrelevante Aspekte herauszulesen. Gefahrenwahrnehmung als individuelle Fähigkeit gilt als grundlegend für sicheres Verhalten im Straßenverkehr. Nach Nagayama (1978) können 50 % aller Verkehrsunfälle auf eine fehlende oder verspätete Wahrnehmung von relevanten Hinweisreizen zurückgeführt werden. Auch laut Rumer (1990) ist mangelnde Gefahrenwahrnehmung, die er als *late detection* bezeichnet, ausschlaggebend für die Mehrzahl aller Unfälle im Verkehr.

Die Gefahrenwahrnehmung erfüllt damit die Kriterien um nach Horswill und McKenna (2004) als eine quantifizierbare Komponente der Fahrexpertise zu gelten, denn Personen, die häufig in der Vergangenheit in Unfälle verwickelt waren, erbringen in Gefahrenwahrnehmungsaufgaben schlechtere Leistungen (z.B. Darby, Murray & Raeside, 2009; McKenna & Horswill, 1999; Pelz & Krupat, 1974; Quimby et al., 1986; Quimby & Watts, 1981; Wells, Tong, Sexton, Grayson & Jones, 2008) und Fahrexperten übertreffen Fahrnovizen darin (z.B. McKenna & Crick, 1991; McKenna & Crick, 1994b; Smith, Horswill, Chambers & Wetton, 2009a).

2.3.1 Modelle zur Einordnung der Gefahrenwahrnehmung in den Risikoverhaltensprozess

Bevor der Fahrer auf eine Gefahr angemessen reagieren und einen Unfall vermeiden kann, muss er einen Reiz zunächst wahrnehmen, erkennen und auch als potenziellen Risikofaktor bewerten. In der Literatur wird *Hazard Perception* zumeist als Prozess, potenziell gefährliche Situationen im Verkehrsgeschehen zu identifizieren, definiert (vgl. McKenna & Crick, 1994; Benda & Hoyos, 1983). Dieser Prozess kann als Teilprozess innerhalb des Risikoverhaltensprozesses betrachtet werden. Deery (1999) schlägt ein Modell vor, das die Prozesse, die dem Fahrverhalten in Reaktion auf potenzielle Gefahren unterliegen, darstellt.

In Abbildung 2.6 werden die wichtigsten Elemente dargestellt, die das Verhalten in Gefahrensituationen beeinflussen. Dieses Modell wurde vor dem Hintergrund entwickelt, dass beim Risikoverhalten perzeptuelle und kognitive Prozesse höherer Ordnung eine Rolle spielen. Diese entwickeln sich laut Deery (1999) relativ langsam, während Novizen basale Fahrfertigkeiten relativ schnell – bereits etwa nach 15 Fahrstunden – erlernen (Hall & West, 1996).

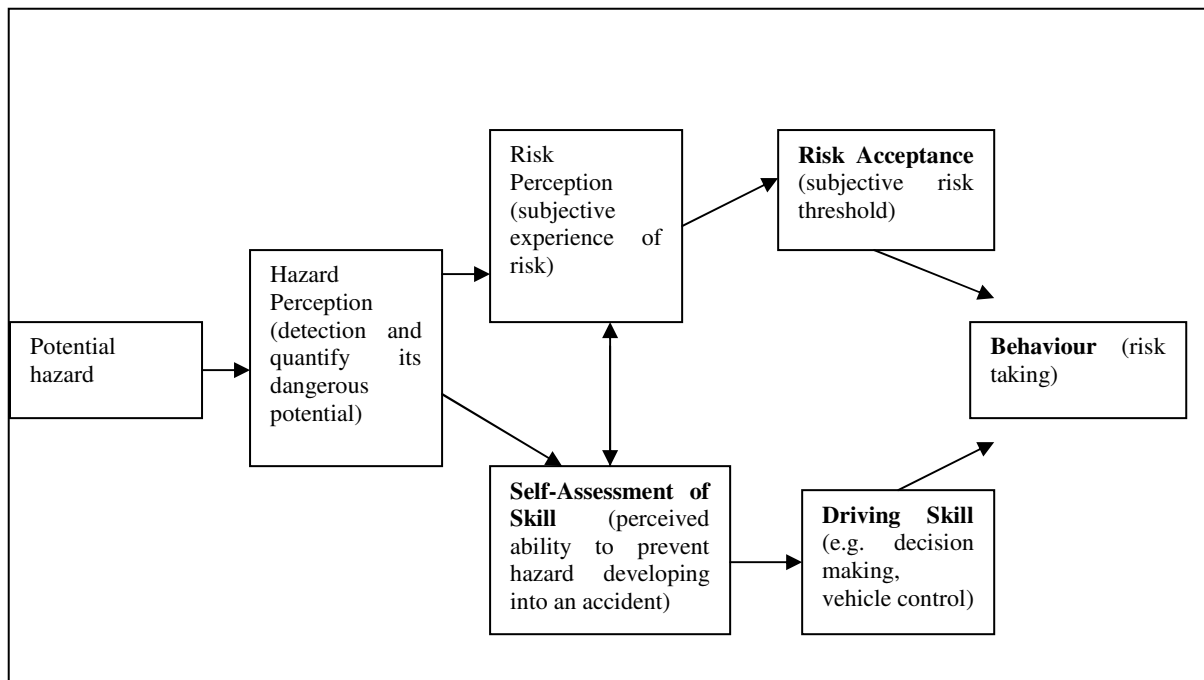


Abbildung 2.6

Modell des Fahrverhaltens als Reaktion auf eine potenzielle Gefahr (nach Deery, 1999)

Deerys Modell zufolge muss ein potenziell gefährlicher Reiz in der Umwelt des Fahrzeuges zunächst erkannt und als potenziell gefährlich eingestuft werden (*Hazard Perception*). Anschließend erfolgt die Einschätzung des Risikos, das von dem Gefahrenreiz ausgeht (*Risk Perception*). Diese Einschätzung des Risikos wird nicht nur von Informationen über die potenziellen Gefahren in der Verkehrsumwelt beeinflusst, sondern auch von der Einschätzung der eigenen Fähigkeiten (*Self-Assessment of Skill*) durch den Fahrer und hat im Gegenzug auch Einfluss auf die Selbsteinschätzung. Die Reaktion auf eine potenzielle Gefahr, die sich im tatsächlichen Verhalten äußert, hängt von den Fertigkeiten bei der Fahrzeughandhabung (*Driving Skill*) und auch von der subjektiven Risikoakzeptanz (*Risk Acceptance*) ab. Die Risikoakzeptanz kann nach Allen und Stein (1987) definiert werden als Level des wahrgenommenen Risikos, das der Fahrer noch bereit ist einzugehen.

Deery (1999) fand in einer kritischen Analyse einschlägiger Literatur Hinweise darauf, dass Unterschiede zwischen Experten und Novizen bei sämtlichen Prozessen, die in das Modell eingehen, gefunden wurden. Novizen waren den Experten in allen Teilprozessen unterlegen. Zudem erwiesen sich alle Prozesse als relevant für die Fahrsicherheit. Besonders im Bezug auf den ersten Teilprozess des Risikoverhaltens, der *Hazard Perception*, gibt es zahlreiche empirische Befunde, dass diese Fähigkeit erfahrungsabhängig ist. Auf diesen Aspekt wird in Kapitel 2.3.3.1 genauer eingegangen.

Grayson et al. (2003) schlagen ein Modell des Risikoverhaltens vor, das einer Definition des Begriffs Gefahrenwahrnehmung entspricht, bei der die angemessene Reaktion auf eine Gefahr eingeschlossen ist. Graysons Modell zufolge sind vier voneinander abgrenzbare Prozesse in das Risikoverhalten einbezogen (vgl. Abbildung 2.7).

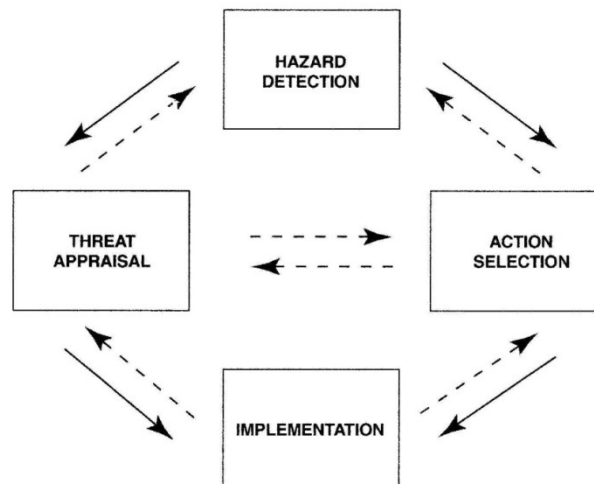


Abbildung 2.7
Modell des Risikoverhaltens (nach Grayson et al., 2003)

Nach dem von Grayson et al. (2003) vorgeschlagenen Modell setzt sich der Prozess des Risikoverhaltens aus den vier aufeinander folgenden Komponenten *Hazard Detection*, *Threat Appraisal*, *Action Selection* und *Implementation* zusammen. Die durchgezogenen Pfeile zeigen mögliche Wege im Risikoverhaltensprozess auf. Dabei sind auch parallele Verarbeitungswege möglich. Die gestrichelten Pfeile stellen mögliche Lernwege dar.

Der erste Teilprozess im Risikoverhaltensprozess wird als *Hazard Detection* bezeichnet und ist als das erste Wahrnehmen eines Reizes zu verstehen. Grayson et al. (2003) nehmen diesen Teilprozess in ihr Modell auf, da sie davon ausgehen, dass das Unfallrisiko deutlich steigt, wenn ein potenzieller Gefahrenreiz übersehen wird. Als Gefahrenreize sind dabei Diskrepanzen zwischen der Wahrnehmung des aktuellen Verkehrsgeschehens und der Vorstellung einer normalen, ungefährlichen Verkehrssituation zu betrachten. Die Identifikation kritischer Reize, die zu einer Gefahr führen könnten, wird in der Regel durch Erfahrung oder Instruktion erlernt, so z.B. die Tatsache, dass zwischen in einem Wohngebiet dicht parkenden Fahrzeugen Kinder hervor laufen könnten. Andere Gefahrenhinweise sind

direkt ersichtlich und es bedarf kaum spezifischer Erfahrung oder vorheriger Anweisung, um sie zu erkennen. Ein Unfall, der unmittelbar vor dem Fahrer von anderen Verkehrsteilnehmern verursacht wird, wird z.B. auch völlig ohne Vorwissen als Gefahr erkannt. Wenn auf dieser Stufe Fehler passieren, ein Gefahrenhinweis also gar nicht wahrgenommen wird, steigt das Unfallrisiko enorm an. Die Komponente *Hazard Detection* ist damit als notwendig für sicheres Verkehrsverhalten anzusehen.

Auch wenn die Wahrnehmung eine bedeutsame Rolle im Risikoverhaltensprozess spielt, haben persönliche Überzeugungen ebenfalls darauf Einfluss. Dieselbe Situation kann von einer Person als gefährlich bewertet werden, wohingegen eine andere Person sie möglicherweise als aufregend empfindet. Hat ein Fahrer eine potenziell gefährliche Situation im Verkehr entdeckt, wird er anschließend bewerten, für wie gefährlich er die Situation hält. Diese zweite Stufe in Graysons Modell wird *Threat Appraisal* genannt. Dabei fließen persönliche Einstellungen und Werte in die Einschätzung einer Verkehrssituation ein. Das Resultat dieser Einschätzung stellt immer die Entscheidung dar, ob eine Handlung erforderlich ist oder nicht. Für jede Person hängt die Bewertung einer Gefahr von überdauernden Persönlichkeitseigenschaften, aktuellem Gemütszustand und Erfahrungswerten ab. Wenn der Fahrer einen wahrgenommenen Reiz als gefährlich interpretiert, wird er angemessen handeln, um einem Unfall vorzubeugen.

Je nachdem wie die Bewertung ausfällt, erfolgt die Auswahl einer als adäquat betrachteten Handlung. Diese Stufe wird im Modell als *Action Selection* bezeichnet. Der Fahrer entscheidet also, wie er auf einen Hinweisreiz reagiert und wählt von seinen zur Verfügung stehenden Handlungsalternativen die passende aus. Diese Auswahl sollte stark damit zusammenhängen, wie automatisiert die einzelnen Handlungsalternativen bisher sind. Auf der letzten Stufe erfolgt die Ausführung der ausgewählten Handlung, die im Modell *Implementation* genannt wird. Selbst wenn der Fahrer einen Gefahrenhinweis entdeckt hat, die Situation als gefährlich einschätzt und das richtige Verhalten wählt, das einen Unfall vermeiden könnte, kann die Unfallvermeidung scheitern. Die Person muss auch über ausreichende Fertigkeiten verfügen, um das gewählte Verhalten – z. B. ein Ausweichmanöver oder die Gefahrenbremsung – korrekt auszuführen. Ob ein Verhalten angemessen ausgeführt werden konnte, beeinflusst auch die Einschätzung zukünftiger Gefahren, weil der Fahrer ein zunehmend genaues Bild seiner eigenen Fähigkeiten erhält. Somit wird er in die Lage versetzt, die Gefahr, die von einer Verkehrsgegebenheit ausgeht, in Abhängigkeit von der eigenen Fähigkeit mit dieser Gefahr erfolgreich umzugehen, zu bewerten.

Das beschriebene Modell des Risikoverhaltens diente als theoretischer Rahmen für ein Forschungsprojekt des Transport Research Laboratory (TRL). Innerhalb dieses Projektes wurde das Modell mittels der Testbatterie CADS (*Computerised Assessment of Driving Skills*), die zu jeder der vier Stufen eine Vielzahl verschiedener Tests beinhaltet, überprüft (Grayson et al., 2003). An der Untersuchung nahmen hauptsächlich junge und unerfahrene Fahrer teil. Die eingesetzten Verfahren waren zum einen neu konzipierte verkehrsspezifische Tests und zum anderen etablierte allgemeine psychometrische Testverfahren und Persönlichkeitsinventare.

Die Untersuchung bestätigte Graysons Modell insofern, als die vier Stufen nur in geringem Maße untereinander korrelierten und somit als unabhängig voneinander gelten können. Die Leistungen in denjenigen Tests, die zu einer Stufe des Risikoverhaltens zählten, korrelierten hingegen hoch. Die Ergebnisse bestätigen somit die angenommene Faktorenstruktur. Zudem wurde die äußere Validität des Modells überprüft. Dazu wurden Zusammenhänge zwischen der Testleistung und der Leistung in einer Fahrprobe ermittelt. Die Leistung in der Fahrprobe auf einer standardisierten Fahrstrecke wurde dabei durch einen ehemaligen Fahrprüfer eingeschätzt. Zusätzlich wurde die gewählte Geschwindigkeit erfasst. Es zeigte sich, dass das Modell insgesamt die Ergebnisse der Fahrprobe bezüglich der Variablen *Gutes Fahren allgemein* und *Tendenz zum Rasen* gut voraussagen konnte.

Wetton et al. (2010) schlagen ein Modell zur Unterteilung der Gefahrenwahrnehmung in drei aufeinanderfolgende Prozesse vor. Die erste Stufe wird analog zum Modell von Grayson et al. (2003) als *Hazard Detection* bezeichnet. Demnach muss eine Person zunächst die Anwesenheit einer potenziellen Gefahr registrieren. Während des darauf folgenden Prozesses muss der Fahrer abschätzen, ob die Bewegungsbahn des potenziellen Gefahrenreizes und/ oder die des eigenen Fahrzeuges das Potenzial hat, eine Kollision zu provozieren. Diese Phase der Antizipation des weiteren Situationsverlaufs wird in Graysons Modell nicht explizit aufgeführt und stellt somit den wesentlichen Unterschied zwischen den beiden Modellen dar. Zuletzt erfolgt die so genannte *Hazard Classification*, bei der der Fahrer entscheiden muss, ob das Ereignis eine evasive Handlung rechtfertigt oder nicht. Jeder der Prozesse wird vom Ergebnis der vorangegangenen Prozesse bedingt. Wetton et al. (2010) vermuten, dass alle drei Prozesse die Reaktionszeiten auf Gefahrenreize beeinflussen. Die Autoren versuchten in zwei Studien die verschiedenen Prozesse, die bei der Gefahrenwahrnehmung beteiligt sind, teilweise zu entkoppeln. Dazu wurde den Teilnehmern ein Test mit statischem Bildmaterial vorgelegt, der lediglich die Geschwindigkeit des *Hazard-*

Detection-Prozesses erfassen sollte (vgl. Kapitel 2.3.4.3) und ein videobasierter Gefahrenwahrnehmungstest, der die Reaktionszeit über den gesamten Prozess der Gefahrenwahrnehmung messen sollte. Im *Hazard-Detection*-Test reagierten unerfahrene Fahrer entgegen den Erwartungen schneller als erfahrene Fahrer. Im *Hazard-Perception*-Test, in dem auf Gefahren in Videos reagiert werden sollte, übertrafen die Experten die Novizen hingegen. Das postulierte Modell konnte mit dem eingesetzten Testmaterial demnach nicht in ausreichendem Maße empirisch validiert werden.

Die verschiedenen dargestellten Modelle zum Risikoverhalten beinhalten als eine erste Stufe immer die einfache Wahrnehmung eines Hinweisreizes. Die darauf folgenden Prozesse sind eine Bewertung des wahrgenommenen Reizes und eine anschließende Verhaltensantwort. Demnach scheint eine erste Wahrnehmung von relevanten Reizen grundlegend für angemessenes Verhalten in Risikosituationen zu sein. Im Folgenden wird daher auf die Bedingungen dafür eingegangen, dass ein relevanter Prozess überhaupt entdeckt wird.

2.3.2 Aufmerksamkeit und das Absuchen der Umwelt in Verkehrssituationen

Aufmerksamkeitsprozesse scheinen beim Fahren eine wichtige Rolle zu spielen, da besonders bei Fahranfängern in vielen Fällen Unaufmerksamkeit zu Unfällen führt (Klauer, S. G. et al., 2006; McKnight & McKnight, 2003). Ein Konzept, das die Gefahrenwahrnehmung und Aufmerksamkeitsprozesse theoretisch verbindet, ist das sogenannte *Situationsbewusstsein* (*Situation Awareness* vgl. Endsley, 1995). Der in der Luftfahrt etablierte Begriff beschreibt dort einen Zustand, der als Voraussetzung für angepasstes Verhalten in einer Umwelt, die sich durch dynamische und interaktive Anforderungen auszeichnet, gelten kann. Dem Konstrukt der *Situation Awareness* wird ein erheblicher Erklärungswert für Vorfälle in der Luftfahrt zugeschrieben (Stelling, 2005). Endsley (1995) beschreibt drei Stufen des Situationsbewusstseins:

1. Wahrnehmung der Elemente der aktuellen Situation
2. Verständnis der aktuellen Situation
3. Antizipation der Entwicklung der Situation

Erst nachdem diese drei Stufen durchlaufen sind, wird eine Entscheidung getroffen und anschließend die entsprechende Handlung ausgeführt. Endsley (1995) geht davon aus, dass sich internale Faktoren wie Ziele, Erwartungen, Fertigkeiten und Fähigkeiten sowie externale Faktoren wie Komplexität der Situation und das Design von Benutzerschnittstellen auf das Situationsbewusstsein auswirken.

Horswill und McKenna (2004) definieren den Zusammenhang zwischen Gefahrenwahrnehmung und Situationsbewusstsein folgendermaßen: „Hazard perception can be considered to be situation awareness for dangerous situations in the traffic environment“ (S. 155). Die Gefahrenwahrnehmung umfasst im Modell des Situationsbewusstseins die beiden ersten Stufen, ein Gefahrenhinweis muss wahrgenommen werden und als kritisch bewertet werden. Ein entscheidender Unterschied besteht darin, dass das Konzept des Situationsbewusstseins generell für die Verarbeitung sämtlicher Reize in der Verkehrsumwelt gilt. Es werden in die Situationsbeurteilung auch Verkehrsregeln und Verkehrszeichen einbezogen. Die Gefahrenwahrnehmung hingegen bezieht sich allein auf das Verhalten in Risikosituationen.

Das Konzept des Situationsbewusstseins, kann auf Verarbeitungsprozesse beim Fahren im Straßenverkehr übertragen werden. Nach Vollrath (2010) muss ein Fahrer, um die Handlungsrelevanz der Situation zu erkennen und weitere Handlungen zu planen, z.B. andere Verkehrsteilnehmer aktiv suchen, sie erkennen und ihr Verhalten antizipieren. Rauch, Gradenegger und Krüger (2009) betonen, dass es insbesondere im Fahrkontext, in dem sich Umweltgegebenheiten kontinuierlich ändern, notwendig ist, das Situationsmodell mit den aktuellen Gegebenheiten abzustimmen.

Zur Messung des Situationsbewusstseins entwickelte Endsley (2000) die *Situation Awareness Global Assessment Technique* (SAGAT). Dabei wird eine Person im Simulator einer realistischen Verkehrssituation ausgesetzt. Gelegentlich wird die Präsentation gestoppt, das Bild verschwindet und der Teilnehmer muss einige Fragen zur zuletzt erlebten Situation beantworten. Ähnlich zu diesem Konzept zur Erfassung des Situationsbewusstseins ist OSCAR aufgebaut (Outil Standardisé pour la Comparaison et l'Analyse des Représentations mentales, Bailly, Bellet & Goupil, 2003). Dabei handelt es sich um ein Verfahren, das überprüfen soll, ob die Fahrer über ein angemessenes mentales Modell der aktuellen Situation verfügen. Bei diesem Verfahren wird anstatt eines schwarzen Bildschirms am Ende eines Videos eine leicht veränderte Situation gezeigt, (z.B. steht an einer Kreuzung eine Ampel, die zuvor nicht vorhanden war). Die Veränderung soll dann durch Anklicken des veränderten

Bereiches angezeigt werden. In einer Studie mit 20 Experten und 20 Novizen konnten die Autoren zeigen, dass Personen mit Fahrerfahrung den Novizen signifikant überlegen waren, sie identifizierten im Durchschnitt fast 20 % mehr Veränderungen (Bailly et al., 2003).

Bei den genannten Verfahren handelt es sich um Befragungsmethoden, die voraussetzen, dass man sich bewusst an Aspekte erinnern und diese auch berichten kann. Implizite Verhaltensweisen, wie sie beim geübten Fahren eine große Rolle spielen sollten, werden nicht berücksichtigt und der Aspekt der Antizipation vernachlässigt (Rauch et al., 2009). Diese Kritikpunkte führten dazu, dass innovative Messmethoden entwickelt wurden. Als fruchtbar erwies sich der Ansatz, zu der Primäraufgabe – dem Fahren im Fahrsimulator – die Beschäftigung mit einer Nebenaufgabe optional anzubieten (Rauch et al., 2009). Die Autoren nahmen an, dass sich Situationsbewusstsein in der, an die aktuelle Verkehrssituation angepasste Ausführung der Nebenaufgabe (Zahlen vorlesen bzw. Interaktion mit einem Fahrerinformationssystem), zeigt. Es konnte festgestellt werden, dass sich die Probanden in vordefinierten kritischen Situationen seltener und eine kürzere Zeitdauer mit den Nebenaufgaben beschäftigten. Die Autoren interpretieren dies als Hinweis auf antizipatorische Vorgänge.

Bei der Navigation durch den Verkehr sollte zudem effektives Absuchen der Umwelt das rechtzeitige Entdecken potenzieller Gefahren und Hinweise auf angemessene Verhaltensweisen in der jeweiligen Verkehrssituation (z.B. Verkehrsschilder) ermöglichen und somit der Gefahrenwahrnehmung dienen. Das *Scanning*-Verhalten einer Person kann mithilfe von Aufzeichnung seiner Blickbewegungen durch einen *Eye Tracker* erfasst werden. Mit der *Eye-Tracking*-Methode kann zuverlässig beobachtet werden, ob der Proband relevante Bereiche des Szenarios fixiert, das er gerade beobachtet (Chapman & Underwood, 1998).

Es konnte gezeigt werden, dass die Unfälle junger Fahrer häufig aufgrund unzureichenden Absuchens der Umwelt zustande kommen (McKnight & McKnight, 2003). In diesem Absuchen der Umwelt nach relevanten Reizen scheinen Experten und Novizen sich zu unterscheiden (z.B. Deery, 1999; Mayhew & Simpson, 1995). Demnach haben Fahranfänger Schwierigkeiten, Verkehrskonstellationen in ihrer gesamten Komplexität wahrzunehmen, da sie ihre Umwelt beim Fahren nur in Ausschnitten wahrnehmen. So verpassen sie, frühzeitig zu erkennen, dass sich Gefahrenpotenzial aufbaut. Novizen fixieren erstens weniger und zweitens eher die statischen Objekte ihrer Fahrumwelt (vgl. Katayama, Motoki, Ochiai & Nakanishi, 1991; Soliday, 1974). Außerdem lenken sie ihre Aufmerksamkeit im Regelfall lediglich auf ein kurzes Stück der Wegstrecke vor ihrem Fahrzeug und schauen beim Fahren

weniger weit in die Ferne als erfahrene Fahrer. Mourant und Rockwell (1972) konnten zeigen, dass Novizen längere Fixationsdauern aufweisen, was darauf hindeutet, dass sie länger brauchten, um Reize, die in den gezeigten Szenarien auftauchten, zu verarbeiten. Außerdem konnten die Autoren beobachten, dass Fahranfänger seltener von einer Seite zur anderen schauen, besonders wenn sie die Fahrspur wechseln.

Erfahrene Fahrer passen ihre Blickbewegungen beim Fahren an die Komplexität der Umgebung an, während Fahrnovizen nicht sensibel auf Komplexitätsvariationen in ihrem Umfeld reagieren (Crundall, Underwood & Chapman, 1998; Underwood, G., 2007). Dies kann eine Ursache dafür sein, dass Fahranfänger längere Latenzzeiten bei der Gefahrenwahrnehmung aufzeigen als erfahrene Autofahrer. Die Tatsache, dass Fahranfänger sich in ihrem *Scanning* nicht an die Verkehrsgegebenheiten anpassen, kann nach Underwood (2007) auf zwei mögliche Ursachen zurückgeführt werden. Erstens kann man vermuten, dass Anfänger noch kein ausgereiftes mentales Modell möglicher Gefahren in bestimmten Verkehrssituationen aufgebaut haben. Zweitens könnten Novizen dem Absuchen ihrer Umwelt nicht in ausreichendem Maße Aufmerksamkeitsressourcen zuwenden, da ein großer Teil davon durch die motorischen Anforderungen des Autofahrens gebunden ist. In vielen Studien, in denen die Gefahrenwahrnehmung durch Computertests gemessen wird, werden an den Prüfling allerdings kaum komplexe motorischen Aufgaben gestellt, die Aufmerksamkeitsressourcen binden. Die Tatsache, dass Experten den Novizen dabei dennoch überlegen sind, spricht für die Hypothese, dass Novizen auf kein ausgereiftes mentales Modell möglicher Gefahren im Verkehr zurückgreifen können.

Chan, Pradhan, Pollatsek, Knodler und Fisher (2010) berichten von einem Simulatorexperiment mit Fahrnovizen und -experten. Dabei hatten die Versuchsteilnehmer zusätzlich zur Fahraufgabe verschiedene Zusatzaufgaben innerhalb des Fahrzeugs (CD-Suche, Kartenlesen, Telefonbedienung) oder außerhalb davon (visuelle Suche auf einem Display auf der Fahrtstrecke) zu erfüllen. Die Augenbewegungen wurden mit einem *Eye Tracker* aufgenommen. Es zeigte sich, dass Novizen mehr Zeit auf Zweitaufgaben innerhalb des Fahrzeuges aufwendeten als Experten.

Aufmerksamkeitsprozesse scheinen im Risikoverhaltensprozess und speziell bei der Gefahrenwahrnehmung in komplexen Verkehrssituationen eine wichtige Rolle zu spielen. Unerfahrene Fahrer scheinen erfahrenen Fahrern dabei unterlegen zu sein. Valide Tests zur Fahrexpertise sollten demnach diesen Aspekt der Gefahrenwahrnehmung beinhalten.

2.3.3 Die Validität von Gefahrenwahrnehmungstests

2.3.3.1 Gefahrenwahrnehmung und Fahrerfahrung

Die Gefahrenwahrnehmung ist eine Teilfertigkeit bei Fahren, die sich durch Erfahrung und Übung entwickeln sollte. Somit sollte auch die Fähigkeit zur Gefahrenwahrnehmung erfahrungsabhängig sein. Das bedeutet, Fahrer, die bereits viel Fahrerfahrung gesammelt haben, sollten sich bei der Gefahrenwahrnehmung anders verhalten als Fahrneulinge. Bei der Leistungsmessung ist zu erwarten, dass Experten die Novizen übertreffen. Zu diesem Ergebnis kamen zahlreiche Studien, die zum einen zeigen konnten, dass Experten schneller auf potenzielle Gefahrenreize reagierten (Biermann, 2007; McKenna & Crick, 1994b; Smith et al., 2009a; Sümer et al., 2007) und zum anderen mehr Gefahren richtig identifizierten (McKenna & Crick, 1994a, 1994b). Im Folgenden werden Studien dargestellt, bei denen Gefahrenwahrnehmungstests im Sinne von verkehrsspezifischen Reaktionstests vorgegeben wurden.

McKenna und Crick (1991) validierten in ihrer Studie einen Test zur Gefahrenwahrnehmung, indem sie diesen Test drei Gruppen von Fahrern vorgaben. Die Teilnehmer wurden den Gruppen anhand ihrer Fahrerfahrung zugeteilt. Es zeigte sich, dass der Test valide zwischen Novizen, erfahrenen Fahrern und Fahrexperten differenzieren konnte. Es zeigte sich ein positiver Zusammenhang zwischen der Fahrerfahrung und der Anzahl der identifizierten Gefahrenhinweise.

Die Autoren führten später eine Studie mit einem simulatorbasierten Gefahrenwahrnehmungstest durch (McKenna & Crick, 1994b). Sie konnten zeigen, dass die Novizen, die ihre Fahrerlaubnis erst seit maximal drei Jahren besaßen, langsamer auf die auftauchenden Gefahren reagierten und auf mehr Gefahren gar nicht reagierten als die Experten, bei denen es sich um Fahrlehrer einer Polizeifahrschule handelte.

Biermann (2007) zeigte in ihrer Studie, in der ein Reaktionstest mit Fotos von Verkehrsszenarien eingesetzt wurde, dass mit zunehmender Fahrerfahrung die Reaktionszeiten auf die gezeigten Bilder mit Gefahrensituationen kürzer wurden. In einer Längsschnittstudie über sechs Monate zeigte sich zudem eine Verbesserung der Gefahrenwahrnehmung, die sich in einer Verkürzung der Reaktionszeiten ausdrückte.

In einer Studie von Sümer et al. (2007) mit dem *Turkish Hazard Perception-Test* wurden Fahrneulingen sowie erfahrenen Fahrern Videos von realen Verkehrsszenarien und computeranimierte Szenarien dargeboten. Es sollte mit Tastendruck immer dann reagiert werden, wenn ein Gefahrenhinweis auftauchte. Es zeigten sich nur für einzelne Szenarien

signifikante Unterschiede zwischen Experten und Novizen, wobei Experten durchgehend schneller reagierten. Außerdem hatte die Studie zum Ergebnis, dass die realen Videos etwas trennschärfer zwischen Novizen und Experten zu differenzieren vermochten als die computeranimierten Versionen. Kritisch ist allerdings anzumerken, dass ein animiertes Szenario nicht jeweils genau einem bestimmten realen Szenario nachempfunden war, was bedeutet, dass auch inhaltliche Aspekte für das Ergebnis ausschlaggebend gewesen sein könnten.

In einem videobasierten *Hazard-Perception*-Test zeigten Smith et al. (2009b), dass die mittleren Reaktionszeiten auf potenzielle Gefahrenhinweise unerfahrener Fahrer höher als die erfahrener Fahrer waren. Außerdem zeigt sich in dieser Studie ein Interaktionseffekt zwischen Fahrexpertise und Schläfrigkeit. Einzelvergleiche zeigten, dass die Novizen langsamer reagierten, wenn der *Hazard-Perception*-Test spät in der Nacht durchgeführt wurde, als wenn sie ihn an einem Vormittag absolvieren mussten. Die Tageszeit hatte jedoch keinen Einfluss auf die Reaktionszeiten der Fahrexperten.

Novizen reagieren jedoch nicht nur langsamer auf Gefahren in gezeigten Verkehrssituationen, sie übersehen zudem besonders häufig verdeckte oder latente Gefahrenhinweise (Borowsky, Shinar & Oron-Gilad, 2007; Pollatsek, Narayanaa, Pradhan & Fisher, 2006; Pradhan et al., 2005).

Hosking, Liu und Bayly (2010) berichten in ihrer Studie ähnliche Ergebnisse für die *Hazard-Perception*-Fähigkeiten von Motorradfahrern. Dabei wurde die Reaktionszeit auf Gefahrenhinweise in simulierten Verkehrsszenarien von drei unterschiedlichen Gruppen verglichen. Es zeigte sich ein signifikanter monotoner Rückgang der Reaktionszeiten von Personen, die sowohl wenig Erfahrung im Auto- als auch im Motorradfahren hatten über Personen, die zwar erfahren im Autofahren aber nicht im Motorradfahren waren, bis zu der Gruppe von Personen, die erfahren in beidem waren.

2.3.3.2 Gefahrenwahrnehmung und Unfallbeteiligung

Ein Schlüsselkriterium zur Bewertung der Relevanz von Messungen, die in Verbindung mit der Sicherheit im Straßenverkehr stehen, stellt ihre Beziehung zur Unfallbeteiligung dar (McKenna & Horswill, 1999). Der Zusammenhang zwischen der Gefahrenwahrnehmungsleistung und der Beteiligung an Unfällen konnte durch zahlreiche Studien belegt werden. Nachweislich erbringen Personen mit wenigen oder keinen Unfällen bessere Leistungen in *Hazard-Perception*-Tests als Personen mit mehr Unfällen (Pelz & Krupat, 1974; Quimby et al., 1986).

Bereits in einer frühen Studie zeigte Currie (1969), dass Personen, die selbst häufiger Unfälle hatten als andere, potenzielle Kollisionen zwischen Modellautos später bemerkten als Personen, die weniger Unfälle hatten. Die unspezifischen Reaktionszeiten, die ebenfalls erfasst wurden, unterschieden sich allerdings nicht zwischen den Gruppen. Dieses Ergebnis ließ vermuten, dass es unabhängig von der basalen Reaktionszeit eine Fähigkeit zu geben scheint, den weiteren Verlauf von Verkehrssituationen zu antizipieren und so bevorstehende Unfälle vorherzusehen. Spicer (1964) hatte bereits 5 Jahre zuvor in einer Studie mit Videomaterial gezeigt, dass Fahranfänger, die seltener in Unfälle verwickelt waren, essentielle Merkmale von Verkehrssituationen besser identifizieren konnten als junge Fahrer, die schon öfter Unfälle hatten.

Pelz und Krupat (1974) führten eine Studie durch, bei der männliche Versuchsteilnehmer innerhalb eines Videos kontinuierlich die aktuelle Gefährlichkeit der gezeigten Situationen einschätzen mussten. Es wurden drei Gruppen von Fahrern verglichen: Fahrer, die bisher nicht an Unfällen beteiligt waren, Fahrer, die bereits an Unfällen beteiligt waren (ohne Verkehrsverstöße) und als letzte Gruppe Fahrer, die bereits Verkehrsverstöße begangen hatten. Unter den Fahrern mit Verkehrsverstößen befanden sich sowohl Personen, mit als auch ohne Unfallbeteiligung in der Vergangenheit. Es zeigte sich, dass sich die drei untersuchten Gruppen bedeutsam in ihrer Reaktion auf Verkehrsgefahren unterschieden. Personen, die bisher keinen Unfall hatten, entdeckten die Gefahrenhinweise früher und reagierten weniger ruckartig als Personen der anderen beiden Gruppen.

Quimby et al. (1986) untersuchten den Zusammenhang zwischen Unfallhäufigkeit und der Zeit, die benötigt wurde, um einen Gefahrenhinweis zu entdecken und darauf zu reagieren. Die Teilnehmer bearbeiteten einen *Hazard-Perception-Test* im Fahrsimulator. Es zeigte sich, dass Personen, die nach eigenen Angaben bisher seltener einen Unfall verursacht hatten, schneller auf Gefahren reagierten als Personen mit mehr Unfällen. Die generelle Reaktionszeit hatte demgegenüber keinen Einfluss auf die Leistung bei der Gefahrenwahrnehmung. Dieses Ergebnis weist darauf hin, dass eine lange Latenzzeit bei der Gefahrenwahrnehmung einen möglichen Risikofaktor darstellt.

Hull und Christie (1993) erprobten in einer Studie ihren *Hazard-Perception-Test*, der für den Einsatz in der Fahrerlaubnisprüfung im Staat Victoria (Australien) konzipiert wurde. Es zeigte sich auch für diesen Test, dass diejenigen unter den über 3000 Studienteilnehmern, die bisher seltener in Verkehrsunfälle verwickelt waren, eine bessere

Gefahrenwahrnehmungsleistung erbrachten als diejenigen Personen, die häufiger an Unfällen beteiligt waren.

Insgesamt hat sich empirisch gezeigt, dass es sich bei den Verfahren zur Messung der Gefahrenwahrnehmung mit Reaktionszeiterfassung um diagnostische Verfahren handelt, die kriteriumsvalide Messungen ermöglichen.

2.3.4 Methoden zur Erfassung der Gefahrenwahrnehmung und der Einsatz neuer Medien

Es hat sich gezeigt, dass Tests zur Gefahrenwahrnehmung Fahrexpertise messen können. Allerdings wurde die Fähigkeit zur Wahrnehmung von Gefahren im Verkehrskontext in verschiedenen Studien auf unterschiedliche Weise erfasst. Zumeist besteht die Aufgabe der getesteten Personen darin, auf Hinweisreize, die auf eine potenzielle Gefahr hindeuten, so schnell wie möglich zu reagieren. Die Verkehrsszenarien, die dem Testteilnehmer präsentiert werden, können entweder durch statisches oder dynamisches Bildmaterial veranschaulicht werden. Bei der Erstellung eines Tests zu Gefahrenwahrnehmung stellt sich daher die Frage, ob man eher bewegte oder unbewegte Bilder darin integrieren sollte. Besonders relevant für die Entscheidung ist dabei, inwieweit sich verschiedene Güteaspekte des Tests in Abhängigkeit von der Präsentationsform des eingesetzten Bildmaterials verändern. Im Verkehrskontext spielt die Bewegung selbst eine bedeutende Rolle. Verkehrsteilnehmer bewegen sich selbst und beobachten andere sich bewegende Verkehrsteilnehmer, was dazu führt, dass visuelle Wahrnehmungseindrücke beim Fahren stetig variieren. Unter Berücksichtigung dieses Aspektes kann angenommen werden, dass durch die dynamische Präsentation von Verkehrsszenarien die ökologische Validität des Tests erhöht wird. Dynamische Präsentationen können allerdings anspruchsvoller für den Rezipienten sein (Mayer, 2008) und erfordern verglichen mit statischen Präsentationen einen höheren finanziellen und zeitlichen Aufwand bei der Erstellung und Präsentation. Aus diesem Grund sollten empirische Forschungsergebnisse für die Entscheidung, welches Präsentationsformat für einen Gefahrenwahrnehmungstest gewählt wird, berücksichtigt werden. Der Einfluss der Dynamik von Darstellungen in *Testaufgaben*, die die Leistung einer Person in einem Inhaltsbereich erfassen, wurde bisher allerdings noch kaum systematisch untersucht. Hingegen war der Vergleich von *Lernmaterial*, das mit statischem vs. dynamischem Bildmaterial angereichert war, schon häufig Gegenstand der Forschung. Daher wird im Folgenden zunächst auf die Ergebnisse der Lehr-Lernforschung zum Einsatz von statischem

vs. dynamischem Bildmaterial beim Lernen eingegangen, um Hinweise für eine optimale Gestaltung von Testaufgaben abzuleiten. Anschließend werden einige Ergebnisse zur Leistungserfassung mithilfe von statischen und dynamischen Präsentationen aus der Domäne Sport vorgestellt, die der Domäne Verkehr insofern ähnelt, als dass sich die meisten Sportarten ebenfalls durch dynamische Vorgänge auszeichnen. Abschließend werden Studien zur Erfassung der Gefahrenwahrnehmung mit statischen und dynamischen Präsentationen dargestellt.

2.3.4.1 Lernen mit statischem vs. dynamischem Bildmaterial

Dynamik als Gestaltungsdimension von Lernmaterial

Aufgabe der Instruktionspsychologie ist es zu ermitteln, wie Lerninhalte optimal präsentiert werden können um das Lernen bestmöglich zu unterstützen. In den letzten 15 Jahren hat der Einsatz von Neuen Medien an Bedeutung für das Lernen gewonnen. Zahlreiche Untersuchungen wurden durchgeführt, um Bedingungen zu identifizieren, unter denen Lerner am meisten von multimedial präsentierten Lerninhalten profitieren (vgl. Mayer, 2001). Dabei wurden Gestaltungsprinzipien auf den vier Dimensionen Kodalität, Modalität, Interaktivität und Dynamik bestimmt, mit deren Hilfe das Lernen erleichtert werden kann (vgl. Mayer, 2001).

Im Folgenden soll die Dimension Dynamik näher betrachtet werden, da diese Arbeit sich inhaltlich mit einer Domäne – dem Autofahren – beschäftigt, die durch dynamische Wahrnehmungseindrücke und Vorgänge gekennzeichnet ist.

Lernmaterialien aber auch Prüfungsaufgaben können rein textuell dargeboten werden oder angereichert mit statischem oder dynamischem Bildmaterial vorliegen. Bei statischem Bildmaterial handelt es sich um unbewegte Bilder, wie Fotos oder Zeichnungen des Lern- bzw. Prüfungsinhaltes. Reales Videomaterial und Computeranimationen werden als dynamisches Bildmaterial bezeichnet. Aufgrund der uneinheitlichen Verwendung der Begriffe *computeranimiert* bzw. *Computeranimation* bedarf es zunächst einer Klärung des Begriffs für die vorliegende Arbeit. Betrancourt und Tversky (2000) definieren ihn folgendermaßen: „computer animation refers to any application which generates a series of frames, so that each frame appears as an alteration of the previous one, and where the sequence of frames is determined either by the designer or the user” (S. 313). Diese Definition schließt jegliche Anwendungen ein, bei denen jedes einzelne Bild einer Abfolge von Bildern die inhaltliche Konsequenz des Bildes darstellt, das ihm vorangegangen ist. Somit werden zum einen

komplett am Computer erstellte bewegte Darstellungen, zum anderen aber auch reale, mit einer Videokamera aufgenommene Filme sowie Abfolgen einzelner inhaltlich aufeinander aufbauender Fotos oder Zeichnungen, in der Definition zusammengefasst. In dieser Arbeit werden allerdings Bildsequenzen in Abhängigkeit von der Art ihrer Erstellung unterschiedlich benannt. „Computeranimation“ bzw. „computeranimiert“ wird im Folgenden nur für Präsentationen verwendet, welche ausschließlich am Computer erstellt wurden. Als „Video“ bzw. „videobasiert“ wird dynamisches Bildmaterial bezeichnet, das durch Abfilmen realer Gegebenheiten entstanden ist.

Einfluss der Dynamik auf das Lernen

Der Einsatz von Animationen, als eine besondere Darstellungsmöglichkeit, die das technologiebasierte Lernen bietet, hat in den letzten Jahren deutlich zugenommen (Lowe & Schnotz, 2008). Tatsächlich wird davon ausgegangen, dass das Verständnis dynamischer Prozesse, das beispielsweise bei Lerninhalten im Bereich der Mechanik relevant ist, durch dynamisch präsentiertes Lernmaterial gefördert wird (Bétrancourt, 2005). Die Ergebnisse der Studien, die den Lernerfolg mit statischem im Gegensatz zu dynamischem Lernmaterial vergleichen, sind uneinheitlich (vgl. Bétrancourt, 2005; Hegarty & Kriz, 2008). So konnten Tversky, Morrison und Bétrancourt (2002) zeigen, dass eine Überlegenheit dynamischen Lernmaterials darauf zurückzuführen war, dass den Lernenden durch die dynamische Präsentation mehr Informationen zu Verfügung gestellt wurden. Wurde die Variable *Informationsmenge* kontrolliert, zeigte sich kein Haupteffekt der Dynamik mehr. Eine Meta-Analyse über 26 Primärstudien, die den Lernerfolg bei Verwendung von statischem bzw. dynamischem Lernmaterial verglichen, zeigte einen moderaten Vorteil dynamischen Lernmaterials gegenüber statischem Lernmaterial (Höffler & Leutner, 2007). Deutlicher wurde das Lernen durch dynamisches Bildmaterial gefördert, wenn es sich bei den animierten Elementen um repräsentative statt um dekorative handelte, wenn also die Animationen tatsächlich Prozesse veranschaulichten, die erlernt werden sollten. Außerdem zeigte sich, dass der positive Effekt der Dynamik sich stärker zeigte, wenn realistisches dynamisches Bildmaterial verwendet wurde. Zuletzt konnten höhere Effekte festgestellt werden, wenn das Lernziel darin bestand, eher prozedurales statt deklaratives Wissen aufzubauen. Mit bewegten Bildern kann demnach nicht unbedingt ein generell besseres Lernergebnis erzielt werden als mit statischen Präsentationen. Inwieweit sich dynamisch präsentiertes Lernmaterial förderlich auf das Lernen auswirkt scheint vom Lerninhalt abzuhängen. Lowe und Schnotz (2008) betonen zudem, dass der lernförderliche Effekt von dynamischen Präsentationen in einem

deutlichen Zusammenhang mit bestimmten Lernermerkmalen steht. So genannte *Aptitude-Treatment-Interaktionen* (kurz ATI), also Wechselwirkungen zwischen Lernermerkmalen und Merkmalen der Gestaltung des Lernmaterials, werden im Folgenden für das Gestaltungsmerkmal Dynamik und die Lernermerkmale Expertise und räumliche Fähigkeiten dargestellt.

Der Einfluss von Expertise auf das Lernen mit dynamischen Darstellungen

Beim Lernen werden die begrenzten Verarbeitungsmöglichkeiten des Arbeitsgedächtnisses belastet. Diese Belastung kann durch ein verbessertes Instruktionsdesign reduziert werden. In Studien, deren Ergebnisse diese Annahme stützen, wurden allerdings in der Regel Personen, die noch über relativ wenig Wissen und Kenntnisse in einem bestimmten Inhaltsbereich verfügten einbezogen. Demgegenüber konnte festgestellt werden, dass die positiven Effekte von gut gestalteten Instruktionsdesigns mit steigendem Vorwissen zunächst verschwinden und sich anschließend sogar umkehren können. Dieser Effekt wird als *expertise-reversal-Effekt* bezeichnet (Kalyuga, Ayres, Chandler & Sweller, 2003). Novizen profitieren demnach von Maßnahmen zu Verbesserung des Instruktionsdesigns eher als Experten. Dieser Effekt wird dadurch begründet, dass das Arbeitsgedächtnis der Personen mit höherem Vorwissen durch diese Gestaltungsmaßnahmen stärker beansprucht wird. Experten hatten bereits die Möglichkeit der adäquaten Schemabildung in einem spezifischen Bereich, während Novizen noch keinerlei Schemata aufbauen konnten. Wird nun das Lernmaterial mit Informationen angereichert, die den Schemaerwerb der Novizen fördern, z.B. durch den Einsatz dynamischer Darstellungen, können diese im Gegenzug bei Experten den Abruf bereits gebildeter Schemata beeinträchtigen. Experten müssen ihre selbst gebildeten Schemata, mit den instruktionell vorgegebenen Informationen abgleichen, dabei müssen sie Redundanzen und Diskrepanzen identifizieren und verarbeiten. Diese Vorgänge binden demnach Ressourcen des Arbeitsgedächtnisses. Kalyuga (2005) konnte zeigen, dass lediglich Personen mit geringem Vorwissen von der guten Gestaltung des Lehrmaterials durch den Einsatz von dynamischen Präsentationen profitierten.

Aus der *cognitive load theory* (Sweller, 1988) wurden bisher allerdings nur Prinzipien für die Gestaltung von geeignetem Instruktionsmaterial für das Lernen von Novizen in der Anfangsphase abgeleitet. Diese Prinzipien sollten dazu führen, dass die Belastung, die durch die Gestaltung des Lernmaterials beim Novizen hervorgerufen wird, reduziert und das Lernen dadurch befördert wird. Die Existenz des *expertise-reversal-Effektes* wird dabei als Indikator dafür angesehen, dass spezifische instruktionale Gestaltungsprinzipien für das Lernen auf

unterschiedlichen Leistungsniveaus, die während des Expertiseerwerbs durchlaufen werden, entwickelt werden sollten (van Gog et al., 2005).

Der Einfluss von räumlichen Fähigkeiten auf das Lernen mit dynamischen Darstellungen

Nicht nur das Vorwissen oder die Expertise können als individuelle Lernermerkmale das Lernen mit Animationen beeinflussen. Auch die räumlichen Fähigkeiten einer Person können sich in Abhängigkeit von der Gestaltung des Lernmaterials auf die Lernleistung auswirken. Das Konzept räumlicher Fähigkeiten, bezieht sich auf die individuelle Fähigkeit einer Person, das visuelle Feld abzusuchen, die Form, Gestalt und Position von Objekten zu verstehen und eine mentale Repräsentation davon aufzubauen und diese auch mental verändern zu können (Carroll, 1993). So genannte *mentale Modelle* sind interne Repräsentationen von Objekten oder Prozessen, die durch den Lerner gedanklich manipuliert werden können. In seiner Vorstellung kann der Lerner ein Objekt beispielsweise um einen bestimmten Winkel drehen, er kann es falten oder zusammenklappen. Dass mentale Modelle innerhalb des Lernprozesses aufgebaut werden, wird als essentiell für das Lernen erachtet (Schnotz, Böckheler & Grondziel, 1999).

Einen zentralen Aspekt räumlicher Fähigkeiten stellt die räumliche Visualisierungsfähigkeit dar (Sternberg, 1990). Mayer und Sims (1994) nehmen an, dass die Fähigkeit zur räumlichen Visualisierung relevant bei der Verarbeitung multimedial präsentierten Materials ist. Vermutlich ist dies der Grund dafür, dass es sich bei der Visualisierungsfähigkeit um den bisher am häufigsten in Studien zur Gestaltung von bildhaftem Lernmaterial gemessenen Faktor räumlicher Fähigkeiten handelt (Höffler, 2010). Testverfahren zur räumlichen Visualisierungsfähigkeit beinhalten zumeist Darstellungen von Objekten, die sich die Probanden in zwei oder drei Dimensionen rotiert oder gefaltet vorstellen sollen.

Einige Aspekte räumlicher Fähigkeiten können sich bei der Informationsverarbeitung von Bildmaterial auf zwei unterschiedliche Arten auswirken (Mayer & Sims, 1994). Einen Unterstützungseffekt guter räumlicher Fähigkeiten sagt die *ability-as-enhancer*-Hypothese voraus. Dieser Hypothese zufolge sollten Personen mit hohen räumlichen Fähigkeiten von bildlich präsentierter Information stärker profitieren als Personen mit geringeren räumlichen Fähigkeiten. Die *ability-as-compensator*-Hypothese hingegen sagt voraus, dass hohe räumliche Fähigkeiten kompensatorisch wirken, wenn eine mangelhafte instruktionale Gestaltung vorliegt, wenn also z.B. statt dynamischer Darstellungen nur statische Präsentationen im Lernmaterial eingesetzt werden. Einige Studien erbrachten empirische

Belege zur Unterstützung der *ability-as-enhancer*-Hypothese (Brünken, Steinbacher & Leutner, 2000; Mayer & Sims, 1994).

Zudem weisen Münzer, Seufert und Brünken (2009) auf den Zusammenhang zwischen der Leistung bei der Verwendung von Lernmaterial, das dynamische Informationspräsentation beinhaltet und den räumlichen Fähigkeiten hin. Eine Meta-Analyse über 19 Primärstudien, in denen der Einfluss räumlicher Fähigkeiten auf das Lernen mit Bildmaterial (statisch und dynamisch) untersucht wurde, deckte insgesamt einen moderaten Vorteil von Lernern mit hohen räumlichen Fähigkeiten auf (Höffler, 2010). Es zeigte sich zudem, dass Lerner mit eher niedrigen räumlichen Fähigkeiten von dynamischen im Gegensatz zu statischen Darstellungen profitieren zu scheinen. Die räumlichen Fähigkeiten standen in einem geringeren Zusammenhang mit den Lernleistungen, wenn zum Lernen Animationen statt statische Darstellungen eingesetzt wurden.

Der seductive-details-Effekt

Lernmaterial kann durch verschiedene Gestaltungsdimensionen beschrieben werden. So kann man es beispielsweise anhand seiner Modalität (z.B. geschriebener vs. gesprochener Text) seiner Kodalität (z.B. Text vs. Bild) oder seiner Dynamik (z.B. Bild vs. Video) einordnen. Diese Gestaltungsdimensionen beziehen sich auf die *äußere* Darstellung von Lernmaterial. Die *innere* bzw. inhaltliche Gestaltung von Texten, Bildern oder Videos gilt aber ebenfalls als relevant für das Lernen.

Dem *seductive-details*-Effekt zufolge sollte es das Lernen negativ beeinflussen, wenn das Lernmaterial zusätzlich zu den relevanten Fakten, mit Informationen angereichert ist, die für den Lerner zwar interessant, aber zugleich zur Lösung der Aufgaben irrelevant sind. Erste Forschungsarbeiten zum *seductive-details*-Effekt bezogen sich auf das Lernen mit Texten (Garner, Gillingham & White, 1989). Später wurde die Definition von *seductive details* auf interessante Details ausgeweitet, die durch andere Medien wie bei Kombinationen aus Text und Bild (Harp & Mayer, 1998) sowie in Animationen (Wright, Milroy & Lickorish, 1999) präsentiert wurden. Es gibt unterschiedliche Auffassungen darüber, wie *seductive details* wirken. Die *Distraction*-Hypothese besagt, dass *seductive details* das Lernen beeinträchtigen, indem das Interesse des Lerners darauf gelenkt und die Aufmerksamkeit vom eigentlichen Lerngegenstand abgelenkt wird (Harp & Mayer, 1998). Nach der *Disruption*-Hypothese wird durch *seductive details* der Aufbau eines kohärenten mentalen Modells unterbrochen. Der *Diversion*-Hypothese zufolge wird hingegen davon ausgegangen, dass der Lerner sich zwar

ein kohärentes mentales Modell bildet, jedoch dafür nicht die wichtigsten Informationen aus dem Lernmaterial extrahiert.

Bisher wurde der *seductive-details*-Effekt lediglich für das Lernen untersucht. Aber auch bei Prüfungsmaterial könnte der Einbau solcher interessanter Details Einfluss auf die Leistung haben. Jarodzka, Scheiter, Gerjets und van Gog (2010) verglichen die Leistungen von Experten (Professoren und Mitarbeiter im Fach Biologie) und Novizen (Biologiestudenten) im Beschreiben von Bewegungsmustern von Fischen. Die Experten übertrafen die Novizen in ihrer Leistung bei der Beschreibung der Bewegungsmuster. Die Experten schauten sich die Videos im Mittel länger an als die Novizen und achteten stärker auf relevante Details, während die Novizen irrelevanten Aspekten in den Videos mehr Beachtung schenkten.

Möglicherweise wird durch die Anreicherung von Prüfungsaufgaben mit *seductive details* die allgemeine Schwierigkeit der Aufgaben erhöht. Denkbar ist auch eine unterschiedliche Wirkung von *seductive details* in Abhängigkeit von der Expertise des Prüflings, da beim Lernen das Vorwissen als Moderator des *seductive-details*-Effekt gilt (Garner, Alexander, Gillingham, Kulikowich & Brown, 1991).

Insgesamt hat sich gezeigt, dass sich die Präsentationsform (statisch vs. dynamisch) des Lernmaterials auf die Lernleistung auswirken kann. Einem eher moderater Haupteffekt, der allgemein für die Vorgabe von dynamisch präsentiertem Lernmaterial spricht, stehen moderierende Faktoren der Lerner selbst (räumliche Fähigkeiten, Expertise) und des Lerninhaltes (deklarativ vs. prozedural; *seductive details*) gegenüber. Diese unterschiedlichen Einflüsse sind theoretisch auch bei der Vorgabe von Testaufgaben, die Bildmaterial in unterschiedlichen Präsentationsformaten enthalten, denkbar.

2.3.4.2 Testen mit statischem vs. dynamischem Bildmaterial

In Kapitel 2.3.4.1 wurde dargestellt, inwieweit dynamische Darstellungen innerhalb des Lernmaterials sich auf das Lernen auswirken können. Wie bereits erwähnt, steht der guten Forschungslage zum *Lernen* mit dynamischem Bildmaterial eher eine geringe Anzahl an Forschungsergebnissen zum Einsatz von dynamischem *Testmaterial* gegenüber. Bei der Erfassung von Fähigkeiten in vielen Sportdomänen lässt sich ein Trend zum Einsatz dynamischer Präsentationen erkennen, die nun zunehmend die in der Vergangenheit häufig verwendeten Bilder im Testmaterial ersetzen. Laut Hodges, Huys und Starkes (2007) wird in sportlichen Domänen besonders die Fähigkeit zum Treffen von angemessenen

Entscheidungen mithilfe von dynamischen Darstellungen überprüft (für Fußball z.B. Ward & Williams, 2003). Während Lowe und Schnotz (2008) anführen, dass die schnelle Ausbreitung von Animationen in Lernmaterial hauptsächlich bedingt durch den enormen technischen Fortschritt auf diesem Gebiet ist und weniger aus inhaltlichen Bedürfnissen resultiert, scheint beim Einsatz dynamischen Bildmaterials bei der Erfassung von Expertise im Sport der umgekehrte Fall vorzuliegen. Eher haben neue und neu formulierte Fragestellungen in diesem Bereich dazu geführt, dass neuartige Methoden zur Erfassung von Expertise entwickelt wurden (Hodges et al., 2007). Die Autoren stellen dar, dass Entscheidungsverhalten im Sport besser mithilfe dynamischer Präsentationen erfasst werden kann als mit statischen, da die ökologische Validität der Animationen höher einzuschätzen ist. Thomas, Gallagher und Lowry (2003) führten eine Metaanalyse über 39 Primärstudien, die in einem Zeitraum von 1987 bis 2002 veröffentlicht wurden, durch. In die Metaanalyse wurden Studien aufgenommen, in denen Experten-Novizen-Vergleiche in der Domäne Sport durchgeführt wurden. Ein wichtiges Ergebnis der Metaanalyse war, dass in Studien mit höherer ökologischer Validität höhere Effektstärken für die Experten-Novizen-Vergleiche festgestellt werden konnten als für die Studien mit niedrigerer ökologischer Validität, besonders wenn die Experten sich auf einem sehr hohen Expertiselevel befanden. Zur Gewährleistung ökologischer Validität werden aber nicht nur dynamische Präsentationen eingesetzt. Hodges, Huys und Starkes (2007) betonen, dass neben dem Präsentationsformat auch Antwortformate weiter entwickelt wurden. Während früher eher geschriebene oder verbale Stellungnahmen erhoben wurden, ging man später dazu über, Reaktionen mit unterschiedlichem Grad an Realitätsnähe zu erfassen.

Das Treffen von Entscheidungen im Sport stellt ähnliche Anforderungen an die Person wie die Entscheidungsfindung beim Fahren, z.B. in Gefahrensituationen. Für den Einsatz von dynamischen Präsentationen in Gefahrenwahrnehmungstests sollten daher ähnliche Gründe sprechen, wie beim Sport, nämlich die damit einhergehende ökologische Validität. Im Folgenden werden zum einen Studien vorgestellt, in denen die Erfassung der Gefahrenwahrnehmung mit statischem Bildmaterial vorgenommen wurde und zum anderen Studien, in denen Gefahrenwahrnehmungstests mit dynamischem Bildmaterial durchgeführt wurden.

2.3.4.3 Methoden der Erfassung der Gefahrenwahrnehmung

Die Erfassung der Gefahrenwahrnehmung mit statischem Bildmaterial erfolgte bisher in empirischen Studien eher selten. Statische Fotos von Verkehrsszenarien kamen z.B. bei

Debus, Leutner, Brünken, Skottke und Biermann (2008) zum Einsatz. Im Rahmen dieser Studie wurde ein Reaktionstest im Sinne eines *Hazard-Perception-Tests* entwickelt. Dabei sollte auf am Computer präsentierte Fotos realer Verkehrssituationen so schnell wie möglich reagiert werden. Die Aufgabe der Probanden bestand darin, bei jedem Foto so schnell wie möglich zu entscheiden, ob die Geschwindigkeit reduziert oder beibehalten werden sollte. Nur im Falle der Geschwindigkeitsreduktion sollte mittels Tastendruck reagiert werden. Eine Besonderheit dieser Studie ist, dass die eingesetzten Fotos inhaltlich anhand der Salienz der gezeigten Gefahrenreize in drei Kategorien eingeteilt wurden: 1. Gefahrenhinweis ohne Gefahrensituation, 2. weniger explizite Gefahrensituation und 3. explizite Gefahrensituation.

Wetton et al. (2010) setzten in einer Studie zur Erfassung der Gefahrenwahrnehmung unter anderem einen Test mit statischem Bildformat ein. Der so genannte *Hazard Change Detection Task* beinhaltet 64 Bildpaare von Verkehrsszenarien. Die beiden Bilder der Bildpaare waren bis auf die An- bzw. Abwesenheit eines Verkehrsteilnehmers der eine potenzielle Gefahr in dem Verkehrsszenario darstellte identisch. Die Bilder eines Bildpaares wurden nacheinander für wenige Millisekunden gezeigt. Die Aufgabe der Versuchsteilnehmer war es, ein Touchscreen, an dem die Bilder präsentiert wurden, an der Stelle zu berühren, wo sich die beiden dargebotenen Bilder unterschieden.

Pelz und Krupat (1974) berichten von einer ersten Videostudie zur Gefahrenwahrnehmung, die von Spicer (1964) durchgeführt wurde. Dabei wurden den Studienteilnehmern Videos gezeigt, die eine Vielzahl verschiedener Verkehrssituationen beinhalteten. Nach jedem Filmsegment hatten die Versuchsteilnehmer die Aufgabe, anhand einer Checkliste einzuschätzen, welche wichtigen Elemente sie innerhalb des Videos entdeckt hatten. Die Besonderheit an dieser Studie war, dass die Filmsegmente relativ lang waren (11 Segmente à 34 Minuten), was bedeutet, dass die Gedächtniskomponente bei dieser Aufgabe eine Rolle spielen sollte.

Pelz und Krupat (1974) zeigten in ihrer Studie einen zehnminütigen Film von Verkehrssituationen in moderatem Stadtverkehr aus der Fahrersicht an einer Leinwand. Innerhalb des Videos tauchten zehn kritische Situationen auf. Einschätzungen bezüglich der Sicherheit dieser Situationen sollten von den Versuchsteilnehmern mit Hilfe eines Einstellhebels am Sitz von Moment zu Moment abgegeben werden. Die Einschätzungen der Versuchsteilnehmer wurden dadurch kontinuierlich registriert.

McKenna und Crick (1991) entwickelten einen Gefahrenwahrnehmungstest, der besonders ressourcensparend war, da er für den standardisierten Einsatz in der englischen

Fahrerlaubnisprüfung konzipiert war. Sie gingen davon aus, dass man bei einem Gefahrenwahrnehmungstest nicht unbedingt sämtliche Aspekte der Fahraufgabe simulieren muss. Auf einem Bildschirm wurden Videos von Verkehrsszenarien dynamisch präsentiert, auf die bei Entdeckung eines Gefahrenhinweises mit der Betätigung einer Taste reagiert werden sollte.

Hull und Christie (1993) setzten ein ähnliches Verfahren ein, bei dem die Probanden jedoch die Entdeckung eines Gefahrenhinweises durch Berühren des Bildschirmes signalisieren sollten.

Smith et al. (2009a) gaben in ihrer Studie einen Test zur Gefahrenwahrnehmung vor, bei dem potenzielle Gefahrenhinweise innerhalb eines am Bildschirm präsentierten Videos durch einen Mausklick lokalisiert werden mussten. Gemessen wurde dabei die Reaktionszeit vom Auftauchen eines kritischen Reizes bis zu seiner richtigen Identifikation per Mausklick.

Wetton et al. (2010) nutzten in einer Studie zur Präsentation von Videos von Verkehrsszenarien einen Touchscreen. Die Aufgabe der Versuchspersonen war dabei, potenzielle Verkehrskonflikte zu antizipieren und sie anzuzeigen, bevor der Fahrer im Video Maßnahmen dagegen ergriff. Im Gegensatz zu den Gefahrenwahrnehmungsaufgaben von Hull und Christie (1993) wurde von den Teilnehmern auch verlangt, den Touchscreen dort zu berühren, wo eine Gefahr zu erwarten war. Nur wenn die Person die Gefahr richtig lokalisierte, wurde ihre Reaktionszeit aufgezeichnet. Die Verkehrskonflikte waren dabei so gewählt, dass es früh genug im Video Hinweisreize auf die Gefahr gab. Bei diesem Format reichte es nicht aus anzuzeigen, wann der Versuchsteilnehmer einen Hinweisreiz entdeckt hatte, sondern er musste auch durch Berührung des Touchscreens anzeigen, wo der Hinweisreiz sich befand.

Es besteht die Möglichkeit, die Gefahrenwahrnehmung mithilfe von statischem oder dynamischem Bildmaterial zu erfassen. Beide Präsentationsformate erscheinen dazu prinzipiell geeignet, wobei das dynamische Format eher der Wahrnehmung und den Anforderungen beim tatsächlichen Fahren zu entsprechen verspricht, was für die ökologische Validität spricht. Ein direkter empirischer Vergleich der beiden Formate steht allerdings noch aus.

2.3.5 Zusammenfassung und Forschungsdesiderate

Die Gefahrenwahrnehmung ist eine wichtige Komponente der Fahrexpertise. Sogenannte *Hazard-Perception-Tests*, die die Reaktion der Probanden auf dargebotene Gefahrensituationen erfassen, dienen der validen Messung der Gefahrenwahrnehmung. Bisher

wurde noch nicht empirisch untersucht, inwieweit sich die Präsentationsform des Bildmaterials (statisch vs. dynamisch) auf die Güte der Tests auswirkt. Dies soll in der vorliegenden Arbeit geleistet werden, um die angemessene Erfassung von Aspekten der Bewältigung von Verkehrssituationen durch ein optimales Instrument zu gewährleisten.

Wie der Erwerb der Gefahrenwahrnehmung als einem Teilaspekt der Fahrexpertise erfolgt und welche Voraussetzungen dafür gegeben sein müssen, ist bisher noch nicht empirisch überprüft worden. Aus den unterschiedlichen Modellen des Expertiseerwerbs, die alle von der Erfahrung als kritischem Einflussfaktor auf die Ausbildung besonderer Fähigkeiten ausgehen, lassen sich verschiedene Erklärungen für die Überlegenheit von Experten gegenüber Novizen bei der Gefahrenwahrnehmung ableiten. Dabei spielen auch Annahmen über den Erwerb dieser Fähigkeit eine Rolle. Es ist möglich, dass die Gefahrenwahrnehmung durch Erfahrung und Übung zu einem automatisierten Vorgang wird. Dafür sprechen insbesondere die mit steigender Fahrerfahrung sinkenden Latenzzeiten bei der Reaktion auf Gefahrenreize. Eine Erklärung dafür könnte sein, dass erfahrene Fahrer weniger kognitive Ressourcen bei der Gefahrenwahrnehmung verbrauchen als unerfahrene Fahrer (Horswill & McKenna, 2004). Ein Modell zum Aufbau von Automatisierung liefert Rasmussens Modell zur Handlungskontrolle (1986). Dabei stellt sich die Frage, ob es sich bei der Gefahrenwahrnehmung ebenfalls um einen Routineprozess handelt, der durch Übung zunehmend automatisiert abläuft. Gegen diese Annahme sprechen laut Horswill und McKenna (2004) verschiedene Studien zu Nebenaufgaben beim Fahren.

Im Sinne der kognitiven Flexibilitätstheorie (Spiro et al., 1988) wäre die Überlegenheit von Experten gegenüber Novizen dadurch zu erklären, dass die Experten bereits reichhaltige Erfahrungen mit Gefahrensituationen und auch mit nicht gefährlichen Situationen aus unterschiedlichen Perspektiven und in unterschiedlichen Kontexten gesammelt haben. Sie sollten daher in der Lage sein, mühelos die aktuelle Situation mit anderen erlebten abzugleichen und entweder Gemeinsamkeiten oder Unterschiede schnell zu erkennen.

Nach dem Fünf-Stufen-Modell des Fertigkeitserwerbs von Dreyfus und Dreyfus würden die schnelleren Reaktionen von erfahrenen Fahrern dadurch erklärt werden, dass Entscheidungen (in diesem Fall: Ist dieser Reiz gefährlich?) beim Experten intuitiv erfolgt. Intuitive Entscheidungen beim Experten benötigen weniger Zeit als die analytischen Entscheidungen der Novizen. Zudem nehmen Experten ihre Umwelt nach dieser Theorie holistisch, also als Ganzes wahr, während Novizen die verschiedenen Elemente eines

Kontextes eher einzeln beachten. Dies sollte es Letzteren erschweren, eine Verkehrssituation in ihrer Komplexität zu überblicken und schnell auf sich entwickelnde Gefahren zu reagieren. All diesen Modellen ist gemeinsam, dass man die Voraussage ableiten kann, dass sich die Gefahrenwahrnehmung mit praktischer Erfahrung verbessert. Inwiefern dies zutrifft kann in einer Längsschnittuntersuchung überprüft werden.

3 Problemstellung

Vor dem Hintergrund der Tatsache, dass junge Fahranfänger ein erhöhtes Risiko haben im Verkehr zu verunglücken, war das Ziel der vorliegenden Arbeit die Untersuchung sicherheitsrelevanter Teilaspekte der Fahrexpertise und ihrer Entwicklung. Als Rahmenwerk für die Erforschung von Expertise in jeglichen Bereichen kann der *expert performance approach* dienen (Ericsson & Smith, 1991a). Demzufolge sollten in einem ersten Stadium repräsentative Aufgabenformate erstellt und hinsichtlich ihrer Validität im Experten-Novizen-Vergleich überprüft werden. Valide Aufgabenformate sollten anschließend eingesetzt werden, um den Erwerb der Expertise innerhalb einer Längsschnittstudie abzubilden und mit bestehenden Modellen des Expertiseerwerbs zu vergleichen. Dieser Schritt entspricht dem letzten Stadium des *expert performance approach*.

Aus dem dargestellten theoretischen und empirischen Hintergrund zu Aspekten von Expertise und Expertiseerwerb allgemein, der Fahrexpertise und ihres Erwerbs im speziellen, sowie der Gefahrenwahrnehmung, als einem relevanten Teilbereich der Fahrexpertise, können verschiedene Ziele und Anforderungen für die vorliegende Arbeit abgeleitet werden.

3.1 Erstellung von Aufgaben zur Erfassung von Teilaspekten der Fahrexpertise

Um im Sinne des ersten Stadiums des *expert performance approach* Aufgaben zu erstellen, mit denen Aspekte von Fahrexpertise valide erfasst werden können, sollte zunächst definiert werden, worin sich Fahrexpertise äußert. Aussagen über die Merkmale von Expertise und Expertiseerwerb in der Domäne Autofahren kann man auf der Grundlage von Befunden und Modellen der Expertiseforschung in anderen Domänen ableiten. Für die Auswahl von Aufgaben, die Aspekte der Ausbildung von Fahrexpertise erfassen sollen, muss zunächst festgelegt werden, welche Kriterien diese Aufgaben erfüllen müssen. Anschließend müssen Inhaltsbereiche für die Aufgaben identifiziert werden, die hinsichtlich der Erfüllung der Kriterien vielversprechend erscheinen. Zudem muss beurteilt werden, welche Entwicklung in den Aufgaben der jeweiligen Inhaltsbereiche vorausgesagt werden kann.

3.1.1 Kriterium Expertise

Experten unterscheiden sich von Novizen in einigen Merkmalen, unabhängig davon, in welcher Domäne die Expertise untersucht wurde (vgl. Kapitel 2.1.2). Für Aufgaben, die

kriteriumsvalide Expertise in einer Domäne erfassen sollen, heißt das, dass sie diese Merkmale im Experten-Novizen-Vergleich abbilden müssen. Experten sollten demnach in den erstellten Aufgaben generell deutlich bessere Leistungen erbringen als Novizen. Da die höhere domänenspezifische Leistungsfähigkeit über sämtliche Domänen hinweg als ein Kriterium für Expertise identifiziert wurde, kann dies auch auf die Domäne Autofahren übertragen werden. Die erstellten Aufgaben müssen demnach dahingehend überprüft werden, ob sie kriteriumsvalide zwischen Experten und Novizen trennen, also ob Experten in ihnen bessere Leistungen erbringen als Personen, die als Novizen bezeichnet werden können. Zur Durchführung solcher Experten-Novizen-Vergleiche ist es notwendig festzulegen, wann eine Person als Experte und wann als Novize gelten kann. Personen als Experten zu klassifizieren, sobald sie einen Führerschein erworben haben, wäre unangemessen, da unerfahrene Fahrer kurz nach dem Erwerb ihrer Fahrerlaubnis noch ein sehr hohes Unfallrisiko aufweisen. Wie bereits in Kapitel 2.1.5 dargestellt wurde, ist die gängige 10-Jahresregel zum Erwerb von Expertise (Chase & Simon, 1973; Ericsson & Crutcher, 1990) als Kriterium für Fahrexpertise zu hoch angesetzt. Etwa nach zwei Jahren Fahrerfahrung wird bereits eine Unfallrate erreicht, die in den darauf folgenden Jahren kaum weiter sinkt. Als Experten werden daher im Folgenden diejenigen Personen betrachtet, die seit über zwei Jahren im Besitz einer Fahrerlaubnis für PKW sind. Das Expertisekriterium wurde bewußt nicht strenger gesetzt (z.B. durch Vorgabe einer Mindest-Kilometerleistung), um hohe Anforderungen an die Sensitivität des Messverfahrens zu stellen.

Alle entwickelten Aufgaben sollen im Sinne der Kriteriumsvalidität die Anforderung erfüllen, dass Personen, die zwei Jahre oder länger eine Fahrerlaubnis der Klasse B besitzen (Experten) bessere Leistungen erbringen als Personen, die noch keine Fahrerlaubnis der Klasse B besitzen (Novizen).

3.1.2 Auswahl von Inhaltsbereichen

Wissenschaftler im Bereich der Verkehrssicherheit verfolgen in der Regel das Ziel, quantifizierbare Komponenten zu identifizieren, die als relevant für Fahrexpertise gelten können (Horswill & McKenna, 2004). Wie in Kapitel 2.2.1.3 dargestellt, sind Aspekte des Anfängerrisikos, die aufgrund mangelnder Erfahrung und Übung im Fahren zu Unfällen führen können, direkt nach dem Fahrerlaubniswerb besonders relevant. Daher sollten valide Aufgaben diese Aspekte abprüfen.

Die Gefahrenwahrnehmung wurde empirisch als eine expertiseabhängige Fähigkeit identifiziert. In Aufgaben zur Gefahrenwahrnehmung erbringen Novizen durchweg

schlechtere Leistungen als Experten (vgl. 2.3.3). Modelle, die die Gefahrenwahrnehmung in den Risikoverhaltensprozess integrieren, zeigen, dass die Gefahrenwahrnehmung sowohl mit Prozessen, die mit dem Anfängerrisiko als auch mit dem Jugendlichkeitsrisiko in Verbindung gebracht werden, zusammenhängt (vgl. 2.3.1). Daher soll die Gefahrenwahrnehmung als ein Inhaltsbereich für Aufgabenformate ausgewählt werden.

In der derzeitigen theoretischen Fahrerlaubnisprüfung wird verkehrsspezifisches Wissen abgefragt. Bisher wurde nicht überprüft, ob dieses Wissen sicherheitsrelevant ist, es sollte jedoch bei der erfolgreichen Bewältigung von Verkehrssituationen notwendig sein. Eine Überprüfung, ob Unterschiede im verkehrsspezifischen Faktenwissen die Unterschiede in der Unfallrate von Experten und Personen, die erst seit Kurzem eine Fahrerlaubnis besitzen, bedingen, steht noch aus. Die Aufgaben der derzeitigen Fahrerlaubnisprüfung, bei denen es sich überwiegend um Mehrfachwahlaufgaben handelt, werden häufig kritisiert, wobei besonders der Inhaltsbereich der Gefahrenlehre als optimierungsbedürftig betrachtet wird (Sturzbecher, Kasper, Bönninger & Rüdell, 2008). Daher sollen auch Wissensaufgaben der derzeitigen Fahrerlaubnisprüfung aus dem Inhaltsbereich der Gefahrenlehre für die vorliegende Arbeit ausgewählt werden.

In der Fahrerlaubnisprüfung werden zudem Aufgaben zu sogenannten „Faustregeln“ bezüglich Anhalteweg, Sicherheitsabstand und Überholweg gestellt. Unfallstatistiken haben gezeigt, dass es sich dabei um relevante Aspekte der Fahrsicherheit handelt (vgl. Kapitel 2.2.1). Daher wird dieser Inhaltsbereich ebenfalls aufgenommen.

Für die Erstellung von Aufgaben zur Optimierung der Fahrerlaubnisprüfung werden demnach die Inhaltsbereiche Gefahrenwahrnehmung, Gefahrenlehre, Abstand und Überholen ausgewählt.

3.1.3 Entwicklung von Expertise in fahr- und verkehrsspezifischen Fertigkeiten

Da im Sinne des dritten Stadiums des *expert performance approach* mit Hilfe von validen Aufgaben die Entwicklung der Expertise untersucht werden soll, müssen Hypothesen dazu abgeleitet werden, wie sich einzelne Komponenten der Fahr- und Verkehrsexpertise ausbilden. Annahmen, die die Aneignung von Expertise betreffen, unterscheiden sich zwischen den unterschiedlichen Domänen. In den meisten Theorien zur Entwicklung von Expertise wird die Möglichkeit zur Erfahrungsbildung in einer Domäne als relevant für den Erwerb von Expertise betrachtet (vgl. Kapitel 2.1.3). Die Tatsache, dass Fahranfänger in

Unfallstatistiken überrepräsentiert sind (z.B. Mayhew et al., 2003), stützt die Annahme, dass Fahrexpertise sich ebenfalls durch Erfahrung und Übung aufbaut. Allerdings hat sich gezeigt, dass es sich bei der Fahrexpertise um ein komplexes Konstrukt mit Teilkomponenten handelt, die sich in unterschiedlicher Geschwindigkeit und auf eine unterschiedliche Weise entwickeln können. Hinweise darauf geben z.B. Studien zur Trennung von Anfänger- und Jugendlichkeitsrisiko (vgl. Kapitel 2.2.1).

Deklaratives Verkehrswissen (Gefahrenlehre und Faustregeln) wird im theoretischen Fahrschulunterricht erworben. Der Expertiseerwerb in diesem Bereich lässt sich am besten durch die *deliberate-practice*-Theorie von Ericsson et al. (1993) beschreiben. Dabei ist in diesem Fall unter Praxis und Übung nicht das praktische Fahren zu verstehen, sondern die gezielte Beschäftigung mit dem Lernmaterial. Wie in der *deliberate-practice*-Theorie vorgesehen, wechseln sich beim Lernen Instruktion (Fahrschulunterricht) und eigenständiges Lernen (Lernen zu Hause mit den Fahrschulbögen, oder Computerprogrammen) ab. Direktes Feedback erhalten die Lernenden ebenfalls, in dem sie ihre Lösungen mit der Musterlösung vergleichen. Aufgaben zum deklarativen Verkehrswissen sollten demnach die Veränderungen während der Fahrausbildung abbilden können.

Expertise bei der Gefahrenwahrnehmung, die sich in zuverlässig schneller Erkennung von Gefahrenreizen zeigt, ist eine Teilkomponente der Fahrexpertise. Die vorgestellten Theorien zum Expertiseerwerb haben gemeinsam, dass zum Aufbau von Expertise Erfahrung und Übung notwendig sind. Rasmussens Modell zufolge, müssten ähnliche Situationen immer wieder wiederholt werden (1986). Spiro et al. (1988) setzen hingegen voraus, dass besonders reichhaltige Erfahrungen gemacht werden. Beides ist nur durch eigenständiges Fahren im Realverkehr möglich. Aufgaben zur Gefahrenwahrnehmung sollten also den Expertiseerwerb, der durch praktische Übung erfolgt, als Leistungsverbesserung abbilden.

3.1.4 Identifikation des angemessenen Präsentationsformates (statisch vs. dynamisch)

In Deutschland wurde die theoretische Fahrerlaubnisprüfung zwischen 2008 und 2010 in ein computerbasiertes Testsystem überführt. Hauptsächlich wurde dabei eine Adaptation der herkömmlichen Aufgaben, die oft eine Zeichnung oder eine Fotografie eines Verkehrsszenarios beinhalten, vorgenommen. Allerdings bietet die computergestützte Testung auch die Möglichkeit des Einsatzes innovativer Präsentationsformate, wie beispielsweise die dynamische Präsentation von Verkehrsszenarien in Form von Videos oder

Computeranimationen. Das Ziel dieser Arbeit ist es zu überprüfen, ob der Einsatz von Dynamik in den Testaufgaben ihre Validität zu steigern vermag. Eine Steigerung der Validität durch die dargebotene Dynamik wäre daran zu erkennen, dass die Experten die Novizen deutlicher in denjenigen Aufgaben übertreffen, die mit dynamischem statt statischem Bildmaterial angereichert sind.

Die Annahme, dass dynamische Präsentationen zu einer Erhöhung des Leistungsunterschieds zwischen den Experten und den Novizen führen könnte, kann durch Ergebnisse der Expertiseforschung begründet werden. Mit steigender Domänenspezifität der Aufgaben wird die Überlegenheit der Experten gegenüber den Novizen deutlicher (Glaser & Chi, 1988). Animierte Verkehrsszenarien sollten spezifischer für die Domäne Autofahren sein, denn beim Fahren handelt es sich um einen dynamischen Vorgang, bei dem sich der Fahrer selbst und auch seine Umgebung in Bewegung befinden. Aufgaben, die dynamische Darstellungen beinhalten, sollten somit besser zwischen erfahrenen und unerfahrenen Fahrern trennen.

Befunde aus Studien zum Lernen mit Animationen weisen darauf hin, dass dynamische Darstellungen zum Verständnis von Lerninhalten beitragen können (Höffler & Leutner, 2007). Aufgaben, die dynamische Darstellungen enthalten, die eher zur Erklärung eines Sachverhaltes dienen, sollten die Aufgaben demnach nur vereinfachen, nicht aber ihre Kriteriumsvalidität erhöhen.

Bisher wurden in der Domäne Autofahren keine Experten-Novizen-Vergleiche mit statischem und dynamischem Testmaterial durchgeführt, weshalb keine eindeutigen Annahmen zum Einfluss von dynamischen Darstellungen in diesem Kontext getroffen werden können. Zudem steht die empirische Untersuchung zum Einfluss statischer vs. dynamischer Darstellungen in *Test*aufgaben allgemein ebenfalls noch aus. Daher werden in dieser Arbeit erste Studien dazu im Kontext der Messung von Fahrexpertise vorgestellt.

Zusammenfassend bestehen die Ziele dieser Arbeit darin, zu untersuchen, durch welche Aufgaben relevante Teilaspekte der Fahrkompetenz erfasst werden können und wie sich die Leistungen darin über die Zeit entwickeln. Dabei werden das verkehrsrelevante Wissen (Gefahrenlehre und Faustregeln) sowie die Gefahrenwahrnehmung fokussiert. Zuvor müssen angemessene Methoden zur Erfassung der einzelnen Aspekte identifiziert werden. Besonders relevant erscheint hierbei eine empirisch gestützte Entscheidung für ein Präsentationsformat, in dem das Bildmaterial zur Begleitung der Aufgaben vorgegeben wird.

3.2 Empirische Vorgehensweise

In zwei Vorstudien werden zunächst die unterschiedlichen Aufgabenformate hinsichtlich ihrer Validität getestet. Die Formate und einzelnen Aufgaben, die sich hinsichtlich ihrer Güte und Differenzierungsfähigkeit zwischen Experten und Novizen bewähren, werden anschließend in einer Längsschnittstudie auf ihre Sensibilität hinsichtlich Instruktion und Training durch Praxis getestet.

Bei beiden Vorstudien werden Experten-Novizen-Vergleiche durchgeführt. Für die vorliegende Arbeit sollen als Novizen diejenigen Personen bezeichnet werden, die bisher noch keine Fahrerlaubnis für PKW erworben haben und somit noch nicht eigenständig (ohne Fahrlehrer) im Verkehr gefahren sind. Personen, die seit über zwei Jahren im Besitz einer Fahrerlaubnis für PKW sind, werden als Experten bezeichnet. Erfahrenen und unerfahrenen Fahrern werden die Aufgaben in den unterschiedlichen Aufgabenformaten entweder in einer statischen oder in einer dynamischen Version vorgelegt. Aus den Ergebnissen der Vorstudien kann zum einen abgeleitet werden, welche Aufgabenformate kriteriumsvalide sind – die Experten übertreffen die Novizen dabei. Zum anderen kann ermittelt werden, welches Präsentationsformat bevorzugt werden soll, da es die Güte der Messung gegenüber dem anderen Präsentationsformat verbessert.

Valide Aufgabenformate gehen im bewährten Präsentationsformat in eine Längsschnittstudie ein. In der Längsschnittstudie werden zu drei Messzeitpunkten ebenfalls Experten und Novizen getestet. Dabei befinden sich die Novizen zum ersten Messzeitpunkt ganz am Anfang ihrer Fahrausbildung. Zum zweiten Messzeitpunkt haben sie ihre Fahrerlaubnisprüfung gerade erfolgreich absolviert. Der dritte Messzeitpunkt findet etwa drei Monate nach dem zweiten Messzeitpunkt statt. Dann konnten die Fahranfänger schon etwas Erfahrung im eigenständigen Fahren sammeln. In Abbildung 3.1 wird die empirische Vorgehensweise veranschaulicht.

Die Längsschnittstudie gibt Auskunft darüber, wie sich die Leistungen der Novizen in den einzelnen Teilaspekten der Fahrexpertise in den verschiedenen Lernphasen entwickeln. Die Experten werden ebenfalls zu drei Messzeitpunkten mit dem Aufgabenmaterial getestet, um mögliche Leistungssteigerungen, die lediglich aufgrund der wiederholten Vorgabe des Testmaterials zustande kommen, zu identifizieren. Da sich die Expertise der Experten innerhalb der Längsschnittstudie kaum weiter entwickeln sollte, ist für diese Gruppe nicht mit einer Leistungssteigerung zu rechnen. Sollten die Experten sich dennoch verbessern, ist dies mit der steigenden Vertrautheit mit dem Testmaterial zu erklären. Die Leistungssteigerung der

Novizen ist demnach nur als Zeichen für Expertiseerwerb zu deuten, wenn sie deutlicher ausfällt als bei den Experten.

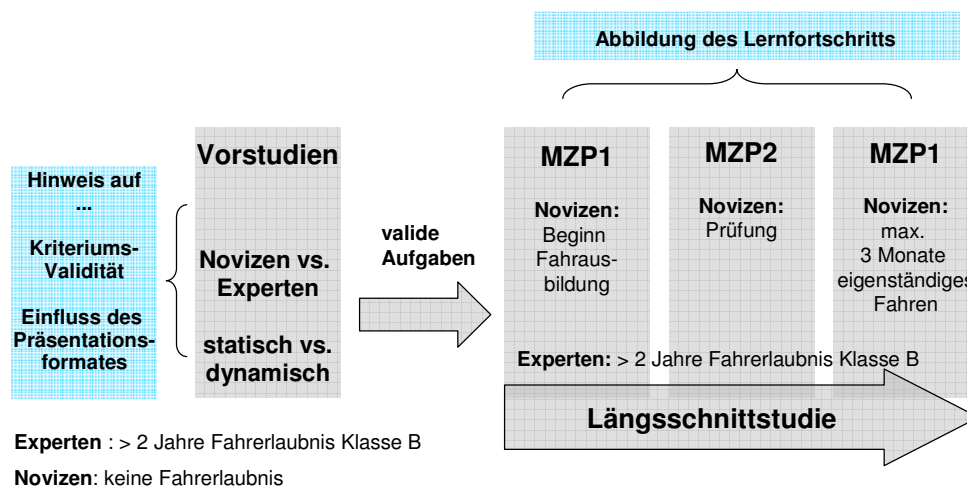


Abbildung 3.1
 Veranschaulichung der empirischen Vorgehensweise

In Abbildung 3.2 werden die Hypothesen für die Längsschnittuntersuchung veranschaulicht.

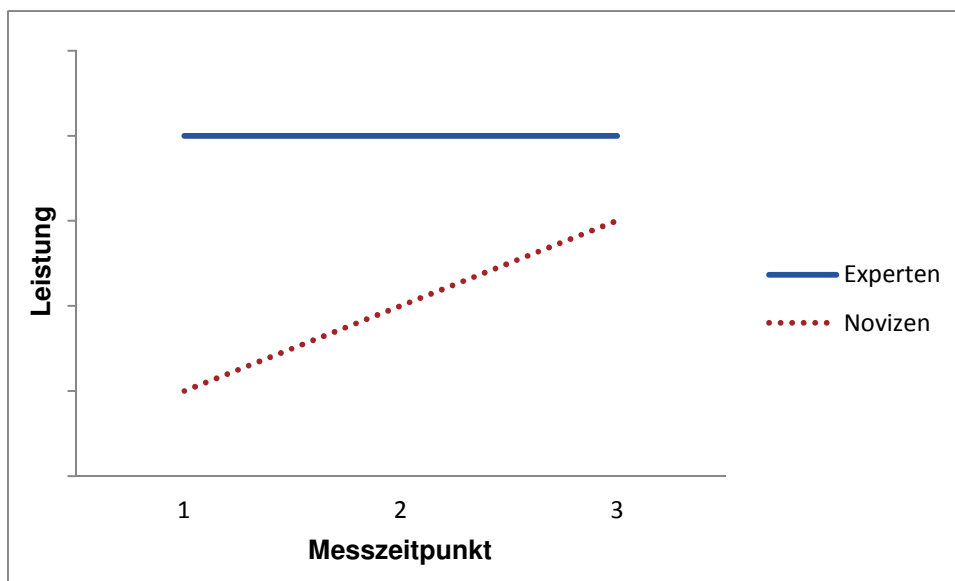


Abbildung 3.2
 Veranschaulichung der Hypothesen für die Längsschnittstudie

4 Vorstudien

Im Folgenden werden zunächst zwei Vorstudien vorgestellt, in denen verschiedene Aufgabenformate zur Erfassung von Teilaspekten der Fahrkompetenz hinsichtlich ihrer psychometrischen Güte überprüft werden. Anschließend wird eine Längsschnittstudie mit drei Messzeitpunkten geschildert, die durch die wiederholte Vorgabe bewährter Aufgabenformate die beginnende Entwicklung von Fahrexpertise abbildet.

Die beiden vorgestellten Vorstudien dienen dazu, geeignete Aufgabenformate zur Erfassung von Teilaspekten der Fahrexpertise zu identifizieren. In beiden Vorstudien werden Experten-Novizen-Vergleiche durchgeführt und das Präsentationsformat des begleitenden Bildmaterials wird experimentell variiert. In Vorstudie 1 werden Aufgaben mit geschlossenem Antwortformat eingesetzt, die deklarative Wissensinhalte zur Gefahrenlehre und die Faustregeln abprüfen. In Vorstudie 2 hingegen werden Aufgaben mit Reaktionszeitantwortformat zur Erfassung der Gefahrenwahrnehmung verwendet.

4.1 Vorstudie 1

4.1.1 Ziele und Fragestellungen

Im Zentrum der ersten Studie stand die Überprüfung der Kriteriumsvalidität von Aufgaben, die deklaratives Verkehrswissen messen sowie den Transfer von so genannten Faustregeln, die in der Fahrschule gelehrt werden. Zudem sollte ermittelt werden, welchen Einfluss die dynamische Gestaltung des Testmaterials auf die Güte der Aufgaben hat.

Eine umfassende Wissensbasis in der jeweiligen Domäne wird als grundlegend für die Expertise angesehen (z.B. Bédard & Chi, 1992). In einigen Modellen des Fertigkeitserwerbs stellt die Phase des Erwerbs von deklarativem domänenspezifischem Faktenwissen die Basis dar, von der aus anschließend erst prozedurale Automatismen ausgebildet werden (vgl. z.B. Anderson, 1982; Fitts & Posner, 1967; Rasmussen, 1986). Es ist daher zu erwarten, dass auch der Erwerb von deklarativem Wissen in der Domäne Autofahren eine wichtige Rolle beim Expertiseerwerb spielt. Häufig wird die Güte der Aufgaben der bestehenden theoretischen Fahrerlaubnisprüfung, die das deklarative Verkehrswissen überprüfen sollen, jedoch in Frage gestellt (z.B. Sturzbecher et al., 2008). In Deutschland wird zurzeit der Einsatz dynamischer Präsentationen von Verkehrsszenarien zur Verbesserung der Fahrerlaubnisprüfung erwogen. Dabei sollen lediglich die momentan eingesetzten statischen Abbildungen durch kurze dynamische Sequenzen ersetzt werden. Der Vorteil der Ergänzung von Aufgaben durch

bewegte Bilder kann darin bestehen, dass die dynamischen Darstellungen der Verkehrssituationen zu einem besseren Verständnis beitragen, wie es aus der Forschung zum Lernen mit Animationen abgeleitet werden kann (Höffler & Leutner, 2007). Ein besseres Verständnis der Aufgaben sollte zu einer verbesserten Reliabilität führen, da Messfehler minimiert werden, die auf mangelndes Instruktionsverständnis zurückgehen. Um zu überprüfen, ob allein die Dynamik Auswirkungen auf die Leistung in den Aufgaben hat, wurden in dieser Studie Aufgaben der theoretischen Fahrerlaubnisprüfung um dynamische Darstellungen der Verkehrsszenarien ergänzt.

Zudem kam in der vorliegenden Studie ein neu entwickeltes Aufgabenformat zum Einsatz. Dabei handelte es sich um Aufgaben, die den Transfer von in der Fahrschule gelernten Faustregeln zu Anhalteweg, Sicherheitsabstand und Überholweg auf bildlich dargestellte Verkehrssituationen verlangten. Auch bei diesem Aufgabenformat, sollte der Einfluss des Präsentationsformats auf die Leistung der Experten und Novizen untersucht werden.

Außerdem sollte geprüft werden, wie räumliche Fähigkeiten mit der Gestaltungsdimension Dynamik interagieren. Da es sich gezeigt hat, dass Aspekte von räumlichen Fähigkeiten sich moderierend auf das Lernen mit unterschiedlichen Präsentationsformaten auswirken (Höffler, 2011), sollte dies auch für das Testen mithilfe von Bildmaterial in unterschiedlichen Präsentationsformaten überprüft werden. Dies dient der Überprüfung der Testfairness und deckt einen möglichen Testbias auf. Solch ein Testbias taucht auf, wenn bestimmte Gruppen durch die Anforderungen verschiedener Aufgaben- oder Präsentationsformate benachteiligt werden und damit der Test anstatt Aspekten der Fahrexpertise eher bestimmte kognitive Fähigkeiten wie z.B. räumliche Fähigkeiten misst. Es handelt sich dabei um einen Aspekt der Validität.

Das Ziel dieser Studie stellte die Überprüfung dar, ob sich zwei Aufgabenformate mit geschlossenem Antwortformat zur Differenzierung von unerfahrenen (Novizen) und erfahrenen Fahrern (Experten) eignen. Des Weiteren sollte untersucht werden, inwieweit die Präsentationsform (dynamisch vs. statisch) die Leistung beeinflusst.

Hypothesen

Wie in Kapitel 2.1 dargestellt, wird Expertise in einer Domäne durch intensive Übung ausgebildet. Durch Übung werden zum einen kognitive Fähigkeiten (spezifisches Wissen, Wahrnehmung und Problemlösestrategien) und zum anderen Können (manuelle Fähigkeiten und Fertigkeiten) erworben. Personen, die schon mehrere Jahre im Besitz eines

Führerscheines sind, sollten durch die zum Teil tägliche Übung im Fahren viel Wissen und Können ausgebildet haben. Fahrschüler hingegen haben bisher nur wenig Erfahrung beim Fahren sammeln können. Wenn die erstellten Aufgaben tatsächlich verkehrsbezogenes Wissen und Können valide überprüfen können, sollten die Experten den Novizen darin überlegen sein.

Hypothese 1: Experten (≥ 2 Jahre Fahrerlaubnis Klasse B) lösen mehr Aufgaben richtig als Novizen (keine Fahrerlaubnis Klasse B). (Haupteffekt Expertise)

Wie in Kapitel 2.3.4.1 dargestellt, konnten Studien, die die Lernleistung bei Verwendung von statischen Bildern der Lernleistung bei Verwendung von bewegten Bildern gegenüber stellen zeigen, dass dynamische Präsentationen unter bestimmten Umständen einen lernförderlichen Einfluss haben können (Höffler & Leutner, 2007). Es wird angenommen, dass die dynamische Darstellung von Lerninhalten, die selbst dynamisch sind, in Anlehnung an die Befunde zum Lernen mit dynamischen Darstellungen, auch bei Testaufgaben verständnisförderlich wirken.

Hypothese 2: Personen, die eine Testversion mit dynamischer Präsentation der Verkehrssituationen bearbeiten, lösen mehr Aufgaben richtig als Personen, die die statische Version bearbeiten. (Haupteffekt Präsentationsformat)

In Kapitel 2.3.4.1 wurde dargelegt, dass die Wirkung von Gestaltungsmerkmalen von persönlichen Voraussetzungen des Lerners abhängen kann. So zeigt sich beispielsweise, dass lediglich Personen mit geringem Vorwissen von nach den Instruktionprinzipien gut gestaltetem Lehrmaterial profitieren (*prior knowledge principle*; Kalyuga, 2005). Die Leistungsfähigkeit von Experten kann dadurch sogar beschränkt werden (*expertise reversal effect*; Kalyuga et al., 2003).

Für die vorliegende Studie sind diese Befunde relevant, da Personen mit hohem und niedrigem Vorwissen und Erfahrung im Bereich Verkehr und Fahren als Experten und Novizen in die Studie einbezogen werden. Es liegt daher die Vermutung nahe, dass Fahrexperten sich eine Verkehrssituation aufgrund ihrer Erfahrung schon anhand einer Beschreibung oder eines statisch dargebotenen Bildes gut vorstellen können. Novizen hingegen könnten von einer dynamischen Darstellung stärker profitieren, da sie die Verkehrssituationen, die in den Aufgaben auftauchen, zumindest als Fahrer selten oder sogar bisher noch nie erlebt haben.

Hypothese 3: Novizen, die die dynamische Testversion bearbeiten, lösen mehr Aufgaben richtig als Novizen denen die statische Version vorliegt. Bei Experten zeigt sich ein geringerer Unterschied. (Interaktion: Präsentationsformat x Expertise).

Das Verursachen von Verkehrsunfällen stellt in vielen Studien ein valides Kriterium für Fahrexpertise dar. Wenn die erstellten Aufgaben Aspekte von Fahrexpertise messen, sollte ihre Leistung darin negativ mit der Anzahl selbst verschuldeter Unfälle zusammenhängen.

Hypothese 4: Es wird ein negativer Zusammenhang zwischen der Menge der richtig gelösten Aufgaben und der Anzahl der durch die Teilnehmer verursachten Unfälle erwartet.

Zur Ergänzung der Hypothesen wurde eine Forschungsfrage aufgestellt, die sich auf den Einfluss individueller kognitiver Fähigkeiten auf die Leistung in den Aufgaben bezieht. In Abschnitt 2.3.4.1 wurde dargestellt, dass verschiedene Präsentationsformate vermutlich unterschiedlich hohe Anforderungen z.B. an die Visualisierungsfähigkeit einer Person stellen können. Hinsichtlich des Zusammenhangs zwischen individuellen räumlichen Fähigkeiten und der Leistung bei der Verarbeitung von Bildmaterial werden zwei mögliche Effekte genannt (Mayer & Sims, 1994). Der *ability-as-enhancer*-Hypothese zufolge profitieren Personen mit hohen räumlichen Fähigkeiten von bildlich präsentierter Information stärker als Personen mit geringeren räumlichen Fähigkeiten. Die *ability-as-compensator*-Hypothese sagt demgegenüber eine kompensatorische Wirkung hoher räumliche Fähigkeiten bei mangelhafter instruktionaler Gestaltung voraus. Individuelle räumliche Visualisierungsfähigkeiten scheinen einen Einfluss darauf zu haben, wie gut eine Person mit unterschiedlichen Präsentationsformaten umgehen kann. Wie sich diese Fähigkeiten exakt auf den Umgang mit statischem vs. dynamischem Bildmaterial auswirken, kann nicht eindeutig vorhergesagt werden.

Forschungsfrage: Besteht ein Zusammenhang zwischen der Visualisierungsfähigkeit und dem Umgang mit unterschiedlichen Präsentationsformaten?

4.1.2 Methode

4.1.2.1 Experimentelles Design und Stichprobe

Der Studie lag ein 2x2-Design mit den Faktoren Expertise (Experten vs. Novizen) und Präsentationsform des Bildmaterials (dynamisch vs. statisch) zugrunde. Als abhängige

Variable wurde die Leistung in den dargebotenen verkehrsspezifischen Aufgaben erhoben. Studien, die *Testaufgaben* mit statischem und dynamischem Testmaterial beinhalten, sind bisher nicht bekannt. Beim *Lernen* mit statischem vs. dynamischem Bildmaterial werden im Durchschnitt mittlere Effekte des Präsentationsformates gefunden (Höffler & Leutner, 2007). Studien, die das fahrspezifische Wissen hinsichtlich der Gefahrenlehre und Faustregeln von Experten mit dem von Novizen vergleichen, sind derzeit nicht bekannt. Experten übertreffen Novizen in domänenspezifischen Aufgaben in der Regel deutlich. Dennoch wird aufgrund unzureichender empirischer Anhaltspunkte zunächst ein mittlerer Effekt des Faktors Expertise angenommen. Die optimale Stichprobengröße wurde nach Faul, Erdfelder, Lang und Buchner (2007) berechnet, wobei das Signifikanzniveau auf 5 % festgelegt wurde und ein mittelhoher Effekt mit einer Wahrscheinlichkeit von mindestens 90 % gefunden werden sollte. Aus diesen Überlegungen resultiert für diese Studie eine optimale Gesamtstichprobengröße von $N = 119$, wobei es für Berechnungen der Haupt- und Interaktionseffekte mithilfe einer Varianzanalyse notwendig ist, auf eine Gleichbesetzung der vier Zellen zu achten (Bortz, 2005).

Insgesamt beteiligten sich 120 Personen an dieser ersten Vorstudie. Die Hälfte der Teilnehmer war männlich. Bei 48 % der Teilnehmer handelte es sich um Novizen. In Tabelle 4.1 wird die Verteilung der Novizen und Experten über die einzelnen Bedingungen dargestellt.

Tabelle 4.1
Verteilung der Versuchsteilnehmer auf die Bedingungen (Vorstudie 1)

N = 120		Präsentationsform des Bildmaterials	
		statisch n = 59	dynamisch n = 61
Expertise	Novizen n = 57	n = 27 (48)	n = 30 (57)
	Experten n = 63	n = 32 (41)	n = 31 (56)

Anmerkung. Die Werte innerhalb der Klammern repräsentieren den Anteil männlicher Probanden in %.

Die Versuchsteilnehmer waren zwischen 16 und 58 Jahren alt ($M = 23.38$; $SD = 7.70$). Dabei waren die Novizen im Durchschnitt 19.47 Jahre alt ($SD = 3.95$) und die Experten 26.92 Jahre ($SD = 8.53$). Ein T-Test für unabhängige Stichproben zeigte, dass damit die Novizen signifikant jünger waren als die Experten ($T(118) = 6.23$; $p < .001$; $d = 1.12$). Der deutliche Altersunterschied musste jedoch in Kauf genommen werden, da Expertise im Autofahren in der Regel positiv mit dem Lebensalter zusammenhängt. Die Altersvarianz innerhalb der beiden Expertisegruppen war allerdings nicht gleich ($F = 8.67$; $p = .004$); das Alter der

Experten variierte stärker. Daher wurde mithilfe von T-Tests für unabhängige Stichproben überprüft, ob sich das Durchschnittsalter der Personen der beiden Expertisegruppen zwischen den Bedingungen (statisch, dynamisch) unterschied. Das durchschnittliche Alter der Novizen betrug für die Teilnehmer mit der statischen Version 19.52 Jahre ($SD = 4.04$) und für diejenigen, die die dynamische Version bearbeitet hatten 19.43 Jahre ($SD = 3.91$). Es zeigte sich kein signifikanter Unterschied ($T(55) = .08$; $p = .94$). Die Experten waren in der statischen Version im Mittel 26 Jahre alt ($SD = 7.15$) und in der dynamischen Version 27.87 Jahre alt ($SD = 9.78$). Die Differenz der Mittelwerte hielt einer Überprüfung auf statistische Signifikanz nicht stand ($T(61) = .87$; $p = .39$). Für beide Expertisegruppen war demnach das Durchschnittsalter der Versuchspersonen, die unterschiedliche Testversionen bearbeiteten, vergleichbar.

Als höchsten Bildungsabschluss, gaben die meisten Personen (57) an, das Abitur bzw. die Fachhochschulreife erworben zu haben. Weitere 36 Personen hatten einen Realschulabschluss, 13 Personen einen Hauptschulabschluss und acht Personen hatten bereits ein Hochschulstudium abgeschlossen. Bei 33 Personen handelte es sich um Schüler. Außerdem nahmen 13 Auszubildende an der Studie teil. Zwölf Personen übten eine Ganztätigkeit aus. Bei 51 Versuchsteilnehmern handelte es sich um Studenten der Universität des Saarlandes.

Im Folgenden werden die beiden Expertisegruppen hinsichtlich ihrer fahrspezifischen Angaben beschrieben. Von den insgesamt 57 Novizen besuchten 37 Personen zum Zeitpunkt der Erhebung gerade die Fahrschule und befanden sich in unterschiedlichen Stadien der Fahrausbildung. Eine weitere Gruppe der Novizen, die aus 11 Personen bestand, gab an, die Absicht zu haben in den nächsten sechs Monaten die Fahrerlaubnis zu erwerben. Die restlichen neun Personen beabsichtigten zurzeit nicht, demnächst einen Führerschein zu erwerben.

Die 63 Experten besaßen seit zwischen zwei und 41 Jahren eine Fahrerlaubnis der Klasse B ($M = 9.43$; $SD = 7.71$). Über Dreiviertel der Personen (76 %) gaben an, über einen eigenen oder geliehenen PKW regelmäßig verfügen zu können. Der größte Anteil der Experten (75 %) gab an, bisher zwischen 2001 und 250 000 Kilometer selbst gefahren zu sein (die Angaben erfolgten in zehn Kategorien unterschiedlicher Größe). Wenige Personen waren bisher weniger als 2000 Kilometer (11 %) oder mehr als 250 000 Kilometer (14 %) selbst gefahren.

4.1.2.2 Versuchsmaterial

Das Versuchsmaterial, das in dieser ersten Vorstudie eingesetzt wurde, bestand aus zwei verschiedenen Arten von Aufgaben zum Verkehrswissen und der Anwendung von Verkehrswissen (s. Anhang B). Die sogenannten „Wissensaufgaben“ sollten dabei deklaratives Verkehrswissen abfragen und die „Einschätzungsaufgaben“ die Fähigkeit zum Einsatz von Faustregeln auf dargebotene Verkehrsszenarien. Des Weiteren wurde ein Testverfahren zur Erfassung eines Aspektes räumlicher Fähigkeiten – der Visualisierungsfähigkeit – integriert (Paper Folding-Test; Ekstrom, French, Harman & Dermen, 1978). Zudem war ein Fragebogenteil enthalten, in dem demographische, sowie Angaben zur Fahrpraxis und Fahrausbildung erfasst wurden. Das gesamte Versuchsmaterial wurde computerbasiert vorgegeben.

Wissensaufgaben

Zur Erfassung deklarativer Wissensinhalte wurden 32 Aufgaben der derzeitigen theoretischen Fahrerlaubnisprüfung (2008) ausgewählt, die aus dem Inhaltsbereich Gefahrenlehre stammen. Davon waren 22 Aufgaben mit einem Bild angereichert, das die zugehörige Verkehrssituation darstellte. Zusätzlich wurden zehn textbasierte Aufgaben einbezogen, die lediglich eine kurze Beschreibung der Verkehrssituation, jedoch kein Bild beinhalteten.

Mithilfe des Computerprogramms VICOM-Editor³ wurde zu jeder Aufgabe, die mit einer realen Fotografie der Verkehrssituation ausgestattet war, ein entsprechendes statisches Bild am Computer erstellt. Abbildung 4.1 zeigt beispielhaft ein Foto, das aus einer Aufgabe der theoretischen Fahrerlaubnisprüfung (Aufgabe Nr. 2.1.08-008) stammt und daneben die am Computer erstellte Entsprechung zu diesem Foto, die in dieser Studie zum Einsatz kam.

Es wurde dabei darauf geachtet, dass die Bilder sich hinsichtlich ihrer wichtigsten Elemente (Art und Anzahl anderer Verkehrsteilnehmer, Wetterbedingungen, Straßenverhältnisse) so ähnlich wie möglich waren. Die am Computer erstellten Bilder der Verkehrsszenarien kamen anstatt der Originalfotos in der statischen Version des Testmaterials zum Einsatz.

³ Bei dem VICOM-Editor handelt es sich um ein Autorenprogramm, mit dem statische und dynamische Verkehrssituationen erstellt werden können. Dieses Programm wurde von der Arbeitsgemeinschaft Technischen Prüfstellen für den Kraftfahrzeugverkehr (TÜV/DEKRA arge tp 21) für das Forschungsprojekt zur Verfügung gestellt.



Abbildung 4.1

Originalbild der Aufgabe 2.1.08-008 (links) und mit dem VICOM-Editor erstelltes Bild (rechts)(aus Malone, Biermann, Buch & Brünken, 2012; S. 36)

Zu jeder Aufgabe wurde ebenfalls mithilfe des VICOM-Editors ein entsprechendes dynamisches Szenario erstellt. Um eine dynamische Version des Originalbildes zu konstruieren, wurde anhand des Bildes und des Originaltextes eine kurze Sequenz erstellt, die darstellte, wie das Fahrzeug auf die Situation, zu der die Frage gestellt wird, zufährt. Jedes Szenario dauerte zwischen neun und zwölf Sekunden. Das Endbild jedes so entstandenen Videos entsprach dem Bild, das in der statischen Version eingesetzt wurde.

Die dynamischen Szenarien beinhalteten keine zusätzlichen Informationen, die notwendig zur Lösung der Aufgabe waren. Dadurch wurde gewährleistet, dass eventuelle Leistungsunterschiede zwischen den beiden Versionen eindeutig auf das Präsentationsformat zurückgeführt werden können. Die Aufgabe der Teilnehmer bestand darin, nachdem sie entweder das Bild zur Aufgabe (statische Version) oder aber das kurze Video zur Aufgabe (dynamische Version) im Vollbildmodus betrachtet hatten, die zugehörige Frage zu beantworten. Es war den Teilnehmern dabei freigestellt, wie oft sie sich das Bild oder Video bildschirmfüllend anschauten. Sie mussten es jedoch mindestens einmal tun, damit die Frage mit den Antwortmöglichkeiten erschien (vgl. Abbildung 4.2).

Die Fragen und Antwortalternativen wurden aus der Originalaufgabe übernommen und waren somit in der statischen und in der dynamischen Version der Aufgabe identisch. Es handelte sich dabei um Mehrfachwahlaufgaben. Von den vorgegebenen Antwortalternativen waren mindestens eine und maximal alle richtig.



Abbildung 4.2

Endbild (Screenshot) einer Beispielaufgabe aus der dynamischen Version des Testmaterials

Einschätzungsaufgaben

Des Weiteren waren 22 Aufgaben zur Einschätzung von Verkehrssituationen und ihrem weiteren Verlauf Teil des Versuchsmaterials. Diese wurden konzipiert, um den Transfer von Faustregeln, die Fahrschüler in der Fahrschule lernen, zu überprüfen. Die Aufgabe der Versuchsteilnehmer war es, für jede einzelne Aufgabe zu entscheiden, ob ein geplantes Fahrmanöver noch sicher ausgeführt werden kann oder nicht. Von jeder Aufgabe wurde eine statische Version mit Bild und eine dynamische Version erstellt, bei der den Teilnehmern die Verkehrssituation in bewegter Form präsentiert wurde.

Unfallstatistiken zeigen auf, dass unangemessene Geschwindigkeit, zu niedriger Abstand und unangebrachtes Überholen zu den Hauptunfallursachen junger Fahranfänger gehören (Statistisches Bundesamt, 2008). Dieser Tatsache wurde dadurch Rechnung getragen, dass drei Aufgabenblöcke zum Einsatz kamen, die die Faustregeln zum Anhalteweg, zum Sicherheitsabstand und zum Überholweg behandeln.

Einer der drei Aufgabenblöcke bezog sich auf die Faustregel zum Anhalteweg. Zur Berechnung des Anhalteweges, der die Summe aus Reaktionsweg und Bremsweg darstellt, wird lediglich die Information über die Geschwindigkeit des eigenen Fahrzeuges benötigt. Der Anhalteweg kann nach Formel (1) berechnet werden.

$$\frac{\text{Geschwindigkeit in km/h}}{10} \times \frac{\text{Geschwindigkeit in km/h}}{10} + \frac{\text{Geschwindigkeit in km/h}}{10} \times 3 \quad (1)$$

Ausgehend von dieser Formel wurden Aufgaben konstruiert, die immer das gleiche Verkehrsszenario aus der Fahrerperspektive darstellten: In einem Wohnviertel mit Einfamilienhäusern läuft ein Kind vor dem PKW auf die Straße. Die Umgebung, also Häuser, Bepflanzung und andere Verkehrsteilnehmer, wurden von Aufgabe zu Aufgabe nur geringfügig verändert (z.B. Passant wurde durch Passantin, blauer VW Golf durch weißen Honda Civic ersetzt usw.). Zwischen den Aufgaben wurden die Parameter Geschwindigkeit des eigenen Fahrzeuges, Abstand zum Kind am Ende des Videos und Witterung variiert. In Tabelle 4.2 wird die Variation der einzelnen Parameter der Aufgaben zum Anhalteweg veranschaulicht.

Abbildung 4.3 und Abbildung 4.4 zeigen die Endbilder der erstellten dynamischen Verkehrsszenarien beziehungsweise die Bilder, die in der statischen Version gezeigt wurden, für zwei Aufgaben mit gleicher Geschwindigkeit aber Variation des Abstandes zum Kind. Die Geschwindigkeit ist dabei auf dem Tachometer im Armaturenbrett abzulesen und wurde zusätzlich im Fragetext jeder Aufgabe erwähnt. Der Abstand zum Kind wurde nicht explizit aufgeführt. Nachdem der Teilnehmer ein Video beziehungsweise ein Bild im Vollbildmodus betrachtet hatte, erschien auf dem Bildschirm zu jeder Aufgabe die Frage: „Können Sie noch rechtzeitig vor dem Kind anhalten?“. Die Teilnehmer konnten dann entweder die Antwortmöglichkeit „ja“ oder „nein“ mithilfe der Computermaus anklicken.

Tabelle 4.2

Variation der Parameter Geschwindigkeit, Abstand und Witterung im Aufgabenblock „Anhalteweg“

Aufgabe	Geschwindigkeit	Abstand	Witterung	Anhalten möglich?
	in km/h	in Metern		
1	50	45	trocken	Ja
2	50	30	trocken	Nein
3	30	30	trocken	Ja
4	30	20	trocken	Ja
5	50	40	Regen und Nässe	Nein
6	30	27	Regen und Nässe	Ja
7	45	44	trocken	Ja
8	30	50	Regen und Nässe	Ja

Anmerkungen. Die Angaben in der zweiten Spalte beziehen sich auf die Geschwindigkeit des Ego-Fahrzeuges. Die Angaben in der dritten Spalte geben den Abstand vom Ego-Fahrzeug zum Kind auf der Straße an.



Abbildung 4.3

Statisches Bild der Einschätzungsaufgabe 1 aus dem Aufgabenblock „Anhalteweg“ (aus Malone et al., 2012; S. 37)



Abbildung 4.4
 Statisches Bild der Einschätzungsaufgabe 2 aus dem Aufgabenblock „Anhalteweg“ (aus Malone et al., 2012; S. 37)

Ein weiterer Aufgabenblock beinhaltete Aufgaben zum adäquaten Sicherheitsabstand zum vorausfahrenden Fahrzeug. Zur Berechnung des angemessenen Sicherheitsabstands beim Fahren auf der Autobahn benötigt man nach der gängigen Faustformel die Geschwindigkeit des eigenen Fahrzeuges. Die Formel (2) zum Sicherheitsabstand lautet:

$$\frac{\text{Geschwindigkeit in km/h}}{2} = \text{Abstand in Metern} \quad (2)$$

Auch in diesem Aufgabenblock wurde bei jeder Aufgabe inhaltlich die gleiche Verkehrskonstellation gezeigt. Das Szenario wird aus der Fahrerperspektive beobachtet. Das Ego-Fahrzeug befindet sich dabei auf der mittleren Spur einer dreispurigen Autobahn und überholt einen LKW, der auf der rechten Spur fährt. Das Ego-Fahrzeug nähert sich einem vorausfahrenden Fahrzeug, das ebenfalls auf der mittleren Spur fährt. Die Frage, die zu dieses Szenarien präsentiert wurde, lautete jeweils: „Halten Sie noch genügend Sicherheitsabstand zum vorausfahrenden PKW ein?“. Variiert wurde zwischen den Aufgaben die Geschwindigkeit des eigenen Fahrzeuges, die Geschwindigkeit des vorausfahrenden Fahrzeuges und der Abstand, der auf dem Bild (statische Version) bzw. am Ende der Animation zwischen den beiden Fahrzeugen vorlag (vgl. Tabelle 4.3).

Die Geschwindigkeit des Ego-Fahrzeuges kann auf dem Tachometer abgelesen werden (vgl. Abbildung 4.5). Die eigene Geschwindigkeit und auch die des vorausfahrenden Pkw's wurden im Fragetext zur jeweiligen Aufgabe angegeben. Der Abstand zwischen den Fahrzeugen wurde den Teilnehmern nicht explizit mitgeteilt.

Tabelle 4.3

Variation der Parameter Geschwindigkeit, Abstand und Witterung im Aufgabenblock „Sicherheitsabstand“

Aufgabe	Geschwindigkeit in km/h	Abstand in Metern	Witterung	Sicherheitsabstand ausreichend?
9	130	75	trocken	Ja
10	130	60	trocken	Nein
11	100	55	trocken	Ja
12	100	40	trocken	Nein
13	130	65	Nebel und Nässe	Nein
14	100	80	Nebel und Nässe	Ja

Anmerkungen. Die Angaben in der zweiten Spalte beziehen sich auf die Geschwindigkeit des Ego-Fahrzeuges. Die Angaben in der dritten Spalte geben den Abstand vom Ego-Fahrzeug zum vorausfahrenden PKW an.



Abbildung 4.5

Statisches Bild der Einschätzungsaufgabe 12 aus dem Aufgabenblock „Sicherheitsabstand“

Weitere acht Aufgaben behandelten die Faustformel zum Überholweg. Um zu berechnen, welcher Wegstrecke es bedarf, um ein vorausfahrendes Fahrzeug zu überholen, benötigt man die eigene Geschwindigkeit (G_{PKW}) in km/h, die Geschwindigkeit des vorausfahrenden

Fahrzeuges (G_{LKW}) in km/h und die Länge beider Fahrzeuge (L) in Metern. Gleichung (3) stellt die gängige Faustformel für den Überholweg dar.

$$\frac{G_{PKW}}{(G_{PKW} - G_{LKW})} \times L = \text{Überholweg in Metern} \quad (3)$$

Der angemessene Abstand zum Gegenverkehr beträgt beim Überholvorgang etwa den doppelten Überholweg. Wie bei den anderen beiden Aufgabenblöcken beinhalteten alle Aufgaben des Aufgabenblocks „Überholweg“ die gleiche Verkehrssituation. Das Ego-Fahrzeug nähert sich auf einer Landstraße von hinten einem LKW, der langsamer fährt als das eigene Fahrzeug. Die Wegstrecke ist gerade und flach, während einige hundert Meter in der Ferne eine enge Kurve zu erkennen ist. Zu diesem Szenario musste die folgende Frage beantwortet werden: „Können Sie, wenn Sie ihre aktuelle Geschwindigkeit beibehalten, noch vor der Kurve gefahrlos überholen?“ Die Geschwindigkeiten der beiden Fahrzeuge sowie der Abstand zur Kurve wurden variiert (vgl. Tabelle 4.4).

Tabelle 4.4

Variation der Parameter Geschwindigkeit des Ego-Fahrzeuges (PKW), Geschwindigkeit des überholten Fahrzeuges (LKW), Abstand und Witterung im Aufgabenblock „Sicherheitsabstand“

Aufgabe	Geschwindigkeit (PKW) in km/h	Geschwindigkeit (LKW) in km/h	Abstand in Metern	Überholweg ausreichend?
15	100 km/h	70 km/h	600	Ja
16	100 km/h	85 km/h	600	Nein
17	85 km/h	70 km/h	600	Nein
18	90 km/h	60 km/h	450	Ja
19	90 km/h	80 km/h	425	Nein
20	75 km/h	60 km/h	450	Nein
21	95 km/h	65 km/h	627	Ja
22	100 km/h	60 km/h	600	Ja

Anmerkung. Die Angaben in der vierten Spalte beziehen sich auf den Abstand vom Ego-Fahrzeug zur Kurve in der Ferne vor der Überholvorgang abgeschlossen sein muss.


Die Länge der Fahrzeuge wurde konstant gehalten, den Versuchsteilnehmern jedoch nicht vorgegeben. Abbildung 4.6 beinhaltet die statische Abbildung aus einer Aufgabe zum Überholweg. Nachdem eine Einschätzungsaufgabe beantwortet wurde, musste der Teilnehmer

auf einer vierstufigen Skala einschätzen, wie sicher er sich bei seiner Antwort war (vgl. Abbildung 4.7, Skala links unten).



Abbildung 4.6
Statisches Bild der Einschätzungsaufgabe 20 aus dem Aufgabenblock „Überholweg“

Sie haben weiterhin die Möglichkeit das Bild zu vergrößern.
Durch Anklicken können Sie es wieder verkleinern.



Sie fahren mit einer Geschwindigkeit von 50 km/h. Können Sie noch rechtzeitig vor dem Kind bremsen?

Bitte markieren Sie die richtige Antwort.

Ja

Nein

Wie sicher sind Sie sich bei Ihrer Antwort?

sehr sicher eher sicher eher unsicher sehr unsicher

Abbildung 4.7
Einschätzungsaufgabe 1 in der statischen Version mit Aufgabenstellung und Einschätzung der Sicherheit

Da bei den Einschätzungsaufgaben ein dichotomes Aufgabenformat vorliegt, bei dem die Wahrscheinlichkeit die richtige Antwort zu erraten bei 50% liegt, wurde eine weitere Differenzierung der gegebenen Antwort von den Teilnehmern verlangt. Bevor eine Einschätzungsaufgabe beantwortet werden konnte, musste das Bild beziehungsweise die Animation mindestens einmal im Vollbildmodus angeschaut werden.

Die Teilnehmer, die die dynamische Form des Testmaterials bearbeiteten, konnten den Beginn jedes Videos durch einen Mausklick selbst bestimmen.

Paper Folding Test

Zur Erfassung der räumlichen Visualisierungsfähigkeit wurde dem *Kit of Factor-Referenced Cognitive Tests* (Ekstrom et al., 1978) der Paper Folding Test (*VZ Visualization*) entnommen. Der Test wurde ausgewählt, da es sich dabei um ein Standardverfahren handelt, das bereits in zahlreichen Studien zu Interaktionseffekten zwischen Lernereigenschaften und Designmerkmalen von multimedialem Lernmaterial eingesetzt wurde (vgl. z.B. Brünken et al., 2000; Mayer & Sims, 1994). Für die vorliegende Studie wurde eine Computerversion dieses Testverfahrens erstellt. Der Paper Folding Test besteht aus insgesamt 20 Aufgaben, die in zwei Blöcken à zehn Aufgaben dargeboten werden. Die Bearbeitungszeit eines Aufgabenblockes ist auf drei Minuten beschränkt. Pro Aufgabe sind die einzelnen Phasen beim Zusammenfalten eines quadratischen Papierblattes abgebildet. Bei der letzten Abbildung wird angezeigt, an welcher Stelle das gefaltete Papier mit einem spitzen Gegenstand komplett durchstanzt wurde. Nachdem dargestellt wurde, in welcher Weise das Papier gefaltet und wo es durchstanzt wurde, werden rechts von einer senkrechten Linie fünf aufgefaltete Blätter mit Löchern abgebildet. Die fünf Rechtecke auf der rechten Seite der Linie stellen Auswahlmöglichkeiten für den Probanden dar. Seine Aufgabe ist es, zu entscheiden, welche der Antwortmöglichkeiten dem gefalteten und durchstanzten Papier entspricht, nachdem es wieder aufgefaltet wurde.

Anhand der Leistung in diesem Test wird die räumliche Visualisierungsfähigkeit beurteilt. Es wird davon ausgegangen, dass der Proband das Blatt mental entfalten und entscheiden muss, zu welchem Ergebnis der Vorgang geführt hat. Die Leistung eines Testteilnehmers entspricht der Anzahl der richtig gelösten Aufgaben in Relation zu den falsch gelösten Aufgaben innerhalb der insgesamt sechs-minütigen Bearbeitungszeit. Zur Berechnung des Leistungswertes der Personen wird eine Ratekorrektur vorgenommen. Formel (4) stellt dar, wie der Leistungswert berechnet wird. Dabei gilt R = Anzahl der richtig

gelösten Aufgaben, F = Anzahl der falsch gelösten Aufgaben und n = Anzahl der Antwortalternativen für jede Aufgabe.

$$L = R - \frac{F}{(n - 1)} \quad (4)$$

Fragebogen

In einem computerbasierten Fragebogen wurden von den Teilnehmern verschiedene Angaben zu ihrer Person erfragt. Zum einen wurden demographische Daten (Geschlecht, Alter, Größe des Wohnortes, Berufstätigkeit, höchster Bildungsabschluss) und zum anderen Angaben zum Führerscheinbesitz (Besitz der Fahrerlaubnis Klasse B und/oder anderer, Stand der Fahrausbildung, gefahrene Kilometer) erfragt. Zudem wurde in diesen Fragebogen eine Erfassung der Erfahrungen mit kritischen Situationen im Verkehr und Unfällen, die bisher verursacht wurden, integriert. Zur Erfassung der kritischen Situationen und Unfälle wurden Fragen eines von einer Projektgruppe um Debus, Leutner und Brünken entwickelten Fragebogens (2008) eingesetzt. Dabei wurden den Teilnehmern zur Erfassung der kritischen Verkehrssituationen 11 konkrete Verkehrssituationen vorgegeben (z.B. „An der roten Ampel beinahe auf ein stehendes Auto aufgefahren“). Danach sollten sie beurteilen, wie oft sie innerhalb der letzten vier Wochen die jeweilige Situation erlebt hatten. Zur Erfassung der von den Teilnehmern verursachten Unfälle wurden die Probanden gebeten, die Anzahl der kleineren Unfälle (ohne Polizeieinsatz), Unfälle mit Sachschaden über 500 Euro und Unfälle mit Personenschaden, die sie in den letzten sechs Monaten und insgesamt seit ihrem Fahrerlaubniswerb verursacht hatten, zu benennen.

4.1.2.3 Durchführung

Zur Anwerbung von Versuchspersonen wurden Aushänge an der Universität des Saarlandes verteilt. Zudem erfolgte eine Direktansprache potenzieller Versuchspersonen im theoretischen Unterricht von insgesamt fünf saarländischen Fahrschulen, die sich zuvor zu einer Kooperation bereit erklärt hatten. Die Versuchspersonen wurden zur Testung entweder in das Computerlabor der Fachrichtung Bildungswissenschaften der Universität des Saarlandes oder in die teilnehmenden saarländischen Fahrschulen eingeladen. Sowohl im Labor als auch in den Fahrschulen wurden Personen mit und ohne Fahrerlaubnis der Klasse B getestet. Die Erhebung wurde in Form von Gruppentestungen durchgeführt, an denen zwei bis maximal zehn Personen gleichzeitig teilnahmen.

Jeder Teilnehmer bearbeitete das gesamte Testmaterial entweder in der statischen oder der dynamischen Version. Die Teilnehmer wurden zufällig auf die Bedingungen des Faktors Präsentationsform (statisch vs. dynamisch) verteilt. Auf den zehn zur Erhebung eingesetzten Laptops war jeweils entweder die dynamische oder die statische Version des Testmaterials eingestellt. Abwechselnd wurde den anwesenden Teilnehmern entweder ein Laptop mit der dynamischen oder der statischen Version zugeteilt. Zudem wurde dabei auf eine Balancierung hinsichtlich des Geschlechtes der Teilnehmer sowie ihrer Expertise geachtet. Hinsichtlich des Faktors Expertise handelte es sich bei der Studie um ein Quasi-Experiment, da den Versuchspersonen nicht zufällig eine Stufe des Faktors Expertise zugeordnet werden konnte, sondern bestehende Gruppen von Novizen (ohne Fahrerlaubnis Klasse B) und Experten (> 2 Jahre Fahrerlaubnis Klasse B) in die Studie einbezogen wurden. Die Erhebungen wurden entweder von der Autorin selbst oder einer studentischen Hilfskraft der Fachrichtung Bildungswissenschaften der Universität des Saarlandes durchführt.

Anschließend an eine kurze Begrüßung der Teilnehmer wurde jedem Teilnehmer eine Aufwandsentschädigung von fünf Euro ausgezahlt. Psychologiestudenten wurden statt einer finanziellen Entlohnung 60 Versuchspersonenminuten bescheinigt. Die Entlohnung erfolgte vor der eigentlichen Testdurchführung, da mit deutlichen Unterschieden bei der Bearbeitungszeit zu rechnen war und Personen, die den Test erst später abschlossen nicht gestört werden sollten. Die Bearbeitungszeiten der unterschiedlichen Versionen unterschieden sich um durchschnittlich etwa zehn Minuten, dabei dauerte die Bearbeitung der statischen Version zwischen 40 und 50 Minuten und die der dynamischen Version 50 bis 60 Minuten. Das komplette Testmaterial wurde am Computer bearbeitet. Dabei wurden sämtliche Instruktionen zu den verschiedenen Testteilen ebenfalls am Computer dargeboten. Verständnisfragen wurden von der jeweiligen Versuchsleiterin beantwortet.

Nach einer kurzen Einführung in Form eines Textes am Bildschirm, der über den Zweck der Untersuchung aufklärte, erfolgte die Bearbeitung des Fragebogens zu persönlichen Angaben, die etwa fünf bis zehn Minuten beanspruchte. Danach wurden den Teilnehmern die Instruktion des Paper Folding Tests sowie anschließend dieser Test selbst vorgelegt. Hiernach bearbeiteten die Teilnehmer die Wissensaufgaben und Einschätzungsaufgaben, denen jeweils eine kurze Instruktion mit einer Beispielaufgabe voranging. Ob einer Person zunächst alle Wissensaufgaben und anschließend alle Einschätzungsaufgaben vorgegeben wurden, oder in umgekehrter Reihenfolge, wurde vom Testprogramm zufällig für jede Person festgelegt. Auch die Reihenfolge der einzelnen Aufgaben innerhalb der Wissens- und der

Einschätzungsaufgaben war randomisiert. Das Ende der Erhebung wurde den Teilnehmern durch eine Mitteilung am Bildschirm signalisiert, nachdem sie die letzte Aufgabe bearbeitet hatten.

4.1.3 Ergebnisse

Für alle statistischen Tests, die im Folgenden berichtet werden, wurde ein Alphaniveau von .05 festgesetzt.

4.1.3.1 Item- und Skalenstatistiken der abhängigen Variablen

Zur Aufgabenselektion und Überprüfung der Reliabilität der drei eingesetzten Aufgabenarten wurden Skalen- und Itemkennwerte berechnet. Die Trennschärfen der einzelnen Aufgaben wurden durch die korrigierte Item-Skala-Korrelation definiert. Neben diesem Trennschärfeindex wurden zur Bestimmung der Fremdtrennschärfe punktbiserale Korrelationen zwischen den einzelnen Items und dem Kriterium Expertise (Experten vs. Novizen) berechnet. Dabei weisen positive Korrelationen darauf hin, dass eine hohe Leistung im jeweiligen Item mit hoher Expertise einhergeht. In Anhang A werden die Kennwerte der einzelnen Aufgaben getrennt berichtet. In Tabelle 4.5 wird ein Überblick über die Itemkennwerte der Wissensaufgaben mit Bildmaterial gegeben.

Die Reliabilität der Gesamtskala der Wissensaufgaben, die statisches oder dynamisches Bildmaterial enthielten, kann mit einer internen Konsistenz von $\alpha = .75$ (nach Kudar Richardson Formula 20) als befriedigend bezeichnet werden. Beide Versionen (statisch und dynamisch) der Skala getrennt weisen mit einer internen Konsistenz von ebenfalls $\alpha = .75$ eine zufriedenstellende Reliabilität auf und sind somit prinzipiell zur Messung von verkehrsspezifischem Wissen geeignet.

Insgesamt zeigte sich, dass es sich bei den Wissensaufgaben in ihrer statischen sowie in ihrer dynamischen Version um reliable Messungen handelt. Aufgrund der zufriedenstellenden internen Konsistenzen können die Aufgaben zu einer Skala zusammengefasst werden, um das verkehrsspezifische Wissen zu erfassen. Für eine Test-Endform, die in der Längsschnittuntersuchung eingesetzt werden sollte, müssen keine Aufgaben ausgeschlossen werden. Einzelne Aufgaben waren sehr leicht ($p > .80$), diese gehen dennoch in eine Endversion ein, da sie auch zwischen Personen differenzieren sollen, die sich am Beginn ihrer Fahrausbildung befinden und damit noch einen niedrigen Wissensstand haben.

Tabelle 4.5
Itemkennwerte der Wissensaufgaben mit Bildmaterial (Übersicht)

	Gesamt	Statisch	Dynamisch
	N = 120	n = 59	n = 61
Mittlere Schwierigkeit (SD)	.67 (.42)	.65 (.42)	.69 (.42)
Schwierigkeit (Min – Max)	.31 – .97	.29 – .95	.33 – .98
Mittlere Trennschärfe	.30	.30	.30
Trennschärfe (Min - Max)	.09 – .53	-.04 – .61	.08 – .53
Mittlere Fremdtrennschärfe	.04	.11	-.03
Fremdtrennschärfe (Min - Max)	-.20 – .20	-.28 – .55	-.36 – .23

Die Eignung der Daten zur Faktorenanalyse wurde durch das Kaiser-Meyer-Olkin-Kriterium überprüft. Mit einem *MSA (measure of sampling adequacy)* von .68 erfüllen die Daten die Voraussetzung zur Berechnung einer Faktorenanalyse. Es wurde eine Hauptkomponentenanalyse durchgeführt, die die einfaktorielle Struktur der Daten bestätigte. Durch den Faktor wurden 18 % der Gesamtvarianz erklärt.

Die Wissensaufgaben, die weder mit statischem noch mit dynamischem Bildmaterial angereichert waren (reine Textaufgaben), wiesen eine mittlere Schwierigkeit von .63 (*SD* = .46) auf. Das Item mit der höchsten Schwierigkeit wurde von 38 % der Personen richtig gelöst, das Item mit der geringsten Schwierigkeit von 85 %. Die korrigierten Item-Skala-Korrelationen variierten zwischen .02 und .50 (*M* = .31). Es konnten Korrelationen mit dem Kriterium Expertise zwischen -.10 und .25 festgestellt werden (*M* = .09). Eine Aufgabe weist einen geringen Zusammenhang mit der Gesamtskala und einen schwach negativen Zusammenhang mit dem Kriterium auf und geht daher nicht in eine Endversion ein. Die interne Konsistenz der Skala ist mit α von .63 als zufriedenstellend einzuschätzen.

Mit einem *MSA* von .70 erfüllen die Daten nach dem Kaiser-Meyer-Olkin-Kriterium die Voraussetzung zur Berechnung einer Faktorenanalyse. Es wurde eine

Hauptkomponentenanalyse durchgeführt, die die einfaktorielle Struktur der Daten bestätigte. Durch den Faktor werden 26 % der Gesamtvarianz erklärt.

Die Einschätzungsaufgaben, deren Itemkennwerte in Tabelle 4.6 dargestellt werden, sind insgesamt von eher geringer Schwierigkeit. Dabei erreichte keines der eingesetzten Items eine befriedigende Trennschärfe und keine Aufgabe weist einen statistisch bedeutsamen Zusammenhang mit dem Kriterium Expertise auf (zweiseitige Testung, Irrtumswahrscheinlichkeit $p < .05$). Auch bei einer Trennung von statischer und dynamischer Version der Testmaterialien nahmen die Itemkennwerte keine befriedigenden Werte an.

Tabelle 4.6

Itemkennwerte der Einschätzungsaufgaben (Übersicht)

	Gesamt N = 120	Statisch n = 59	Dynamisch n = 61
Mittlere Schwierigkeit (SD)	.77 (.43)	.71 (.44)	.73 (.43)
Schwierigkeit Min – Max	.39 – .91	.39 – .90	.38 – .92
Mittlere Item-Skala-Korrelation	.07	.12	.02
Trennschärfe Min – Max	-.14 – .26	-.24 – .34	-.17 – .23
Mittlere Fremdtrennschärfe	.01	.02	-.01
Fremdtrennschärfe Min – Max	.13 – .23	-.12 – .35**	-.22 – .20

Anmerkung. ** $p < .001$

Die Reliabilität der Gesamtskala ist mit einer internen Konsistenz von $\alpha = .28$ (nach Kudar Richardson Formula 20) als unbefriedigend zu bezeichnen. Weder die statische ($\alpha = .41$) noch die dynamische Version ($\alpha = .09$) für sich genommen sind als konsistente Messungen zu bezeichnen. Auch durch den Ausschluss einzelner Items lassen sich keine deutlich höheren internen Konsistenzen erzielen. Die Einschätzungsaufgaben können demnach nicht zu einer Skala zusammengefasst werden. Da die Einschätzungsaufgaben inhaltlich in drei Blöcke unterteilt werden können, wurde untersucht, ob drei Unterskalen anstatt einer Gesamtskala angemessener sind. Mit einem MSA von .59 erfüllen die Daten nach dem Kaiser-Meyer-Olkin-Kriterium die Voraussetzung zur Berechnung einer Faktorenanalyse. Es wurde eine Hauptkomponentenanalyse durchgeführt, die eine dreifaktorielle Struktur jedoch nicht bestätigen konnte. Bei einer einfaktoriellen Lösung werden lediglich 14 % der Gesamtvarianz erklärt. Trennt man die Daten nach der eingesetzten Testversion (statisch vs. dynamisch), dann zeigt sich, dass nach dem Kaiser-Meyer-Olkin-Kriterium eine Faktorenanalyse mit den

Daten, die mit der statischen Version erhoben wurden ($MSA = .56$), nicht aber mit den Daten aus der dynamischen Version ($MSA = .41$) durchgeführt werden darf. Der Eigenwerteverlauf, der bei einer Analyse der Daten aus der statischen Version resultiert, weist zwar auf eine dreifaktorielle Struktur hin (Varianzaufklärung durch die drei Faktoren 37 %), inhaltlich ist aber keiner der drei Faktoren eindeutig interpretierbar.

Die Korrelationsanalyse zeigt einen geringen positiven Zusammenhang zwischen der Leistung in den Wissensaufgaben mit Bildmaterial und der Leistung in den Einschätzungsaufgaben ($r = .18$; $p = .04$).

Zudem konnten positive Zusammenhänge von mittlerer Höhe sowohl zwischen der Leistung in den Einschätzungsaufgaben und den Textaufgaben ($r = .28$; $p = .002$) als auch zwischen den Wissensaufgaben mit Bildmaterial und den reinen Textaufgaben ($r = .61$; $p < .001$) festgestellt werden.

4.1.3.2 Visualisierungsfähigkeit

Für die Leistung der Probanden im Paper Folding Test, der zur Erfassung der Visualisierungsfähigkeit eingesetzt wurde, wurde ein Leistungswert berechnet (vgl. Kapitel 4.1.2). Insgesamt erreichten die Teilnehmer im Durchschnitt einen Leistungswert von $M = 0.24$ ($SD = 2.0$; $Min = -4.0$; $Max = 4.5$). Zum Vergleich der Leistung der Experten und der Novizen wurde ein T-Test für unabhängige Stichproben gerechnet. Es zeigte sich, dass die Experten ($M = 0.66$; $SD = 1.86$) höhere Leistungen erzielten als die Novizen ($M = -0.22$; $SD = 2.07$). Die Differenz der Mittelwerte war statistisch bedeutsam ($T(120) = 2.45$; $p = .02$; $d = .45$). Aufgrund dieser Gruppenunterschiede geht die Variable Visualisierungsfähigkeit als Kovariate in die Varianzanalysen ein.

Es zeigte sich eine niedrige positive Korrelation zwischen der Leistung im Paper Folding Test und der Leistung in den Wissensaufgaben mit Bildmaterial, die allerdings einer Überprüfung auf statistische Signifikanz nicht standhielt ($r = .10$, $p > .05$). Demgegenüber zeigte sich ein mäßig hoher positiver Zusammenhang zwischen der Visualisierungsfähigkeit und der Leistung in den Einschätzungsaufgaben ($r = .21$, $p = .02$). Außerdem konnte eine mäßig hohe positive Korrelation der Leistung im Paper Folding Test und der Leistung in den reinen Textaufgaben ($r = .32$, $p < .001$) nachgewiesen werden.

4.1.3.3 Antwortsicherheit

Zu jeder Einschätzungsaufgabe wurde erhoben, wie sicher sich die Probanden bei ihrer Antwort waren. In die Analyse der Antwortsicherheiten wurden die Angaben von 113

Teilnehmern (55 Novizen und 58 Experten) einbezogen, die restlichen sieben Teilnehmer hatten keine Angaben zu ihrer Antwortsicherheit gemacht. Insgesamt waren sich die Teilnehmer der Richtigkeit ihrer Antworten im Durchschnitt eher sicher ($M = 1.88$; $SD = 0.41$; $Min = 1.09$; $Max = 3.41$). Die durchschnittliche Antwortsicherheit der Experten ($M = 1.82$; $SD = .46$) und der Novizen ($M = 1.94$; $SD = .35$) unterschieden sich nicht statistisch bedeutsam ($T(113) = 1.65$; $p = .102$).

Die Antwortsicherheit korrelierte mit $r = -.07$ gering positiv mit der Anzahl der richtigen Lösungen in den Einschätzungsaufgaben. Der Zusammenhang erwies sich nicht als statistisch bedeutsam ($p > .05$).

4.1.3.4 Einfluss von Expertise und Präsentationsform auf die Leistung

Ziel der Studie war es, den Einfluss von Fahrexpertise und Präsentationsformat des eingesetzten Bildmaterials auf die Leistung in den Wissensaufgaben und den Einschätzungsaufgaben zu ermitteln. In Tabelle 4.7 sind die durchschnittlichen Leistungen der beiden Expertisegruppen in den beiden Aufgabenformaten zu den unterschiedlichen Bedingungen angegeben.

Tabelle 4.7

Mittelwerte und Standardabweichungen der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch; Maximalpunktzahl je Aufgabenart: 22)

	Experten/ statische Version n = 32		Novizen/ statische Version n = 27		Experten/ dynamische Version n = 30		Novizen/ dynamische Version n = 31	
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>
Wissensaufgaben (Min = 0; Max = 22)	15.34	(2.51)	13.15	(4.66)	14.81	(3.24)	15.60	(4.24)
Einschätzungsaufgaben (Min = 0; Max = 22)	15.56	(2.73)	15.33	(2.15)	15.64	(2.36)	15.23	(2.20)

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung

Zwischen den beiden Expertisegruppen lag ein signifikanter Unterschied in der Kontrollvariable Visualisierungsfähigkeit vor. Daher wurde zur Überprüfung des Einflusses von Expertise und Präsentationsform auf die Leistung eine Kovarianzanalyse mit der Kovariate Visualisierungsfähigkeit gerechnet. Die unabhängigen Variablen waren dabei die Expertise (Experten vs. Novizen) und das Präsentationsformat (dynamisch vs. statisch).

Als abhängige Variable wurde zunächst die Leistung in den Wissensaufgaben in die Kovarianzanalyse einbezogen. Die Wissensaufgaben, die nur Text beinhalteten, wurden aus dieser Analyse ausgeschlossen, da hierbei der Faktor Präsentationsformat nicht variiert wurde. Die Kovarianzanalyse zeigte weder einen signifikanten Haupteffekt der Kovariate Visualisierungsfähigkeit ($F(1, 115) = .56; p = .46$) noch der Expertise ($F(1, 115) = 0.71; p = .40$). Ebenfalls konnte kein Haupteffekt der Präsentationsform ($F(1, 115) = 1.97; p = .16$) gezeigt werden. Die Varianzanalyse zeigte einen Interaktionseffekt der Faktoren Expertise und Präsentationsform ($F(1, 115) = 4.85, p = 0.03, \eta_p^2 = .40$).

Abbildung 4.8 verdeutlicht den Interaktionseffekt. Die Novizen erbrachten in der statischen Version geringere Leistungen als in der dynamischen Version, während es im Falle der Experten eher umgekehrt war.

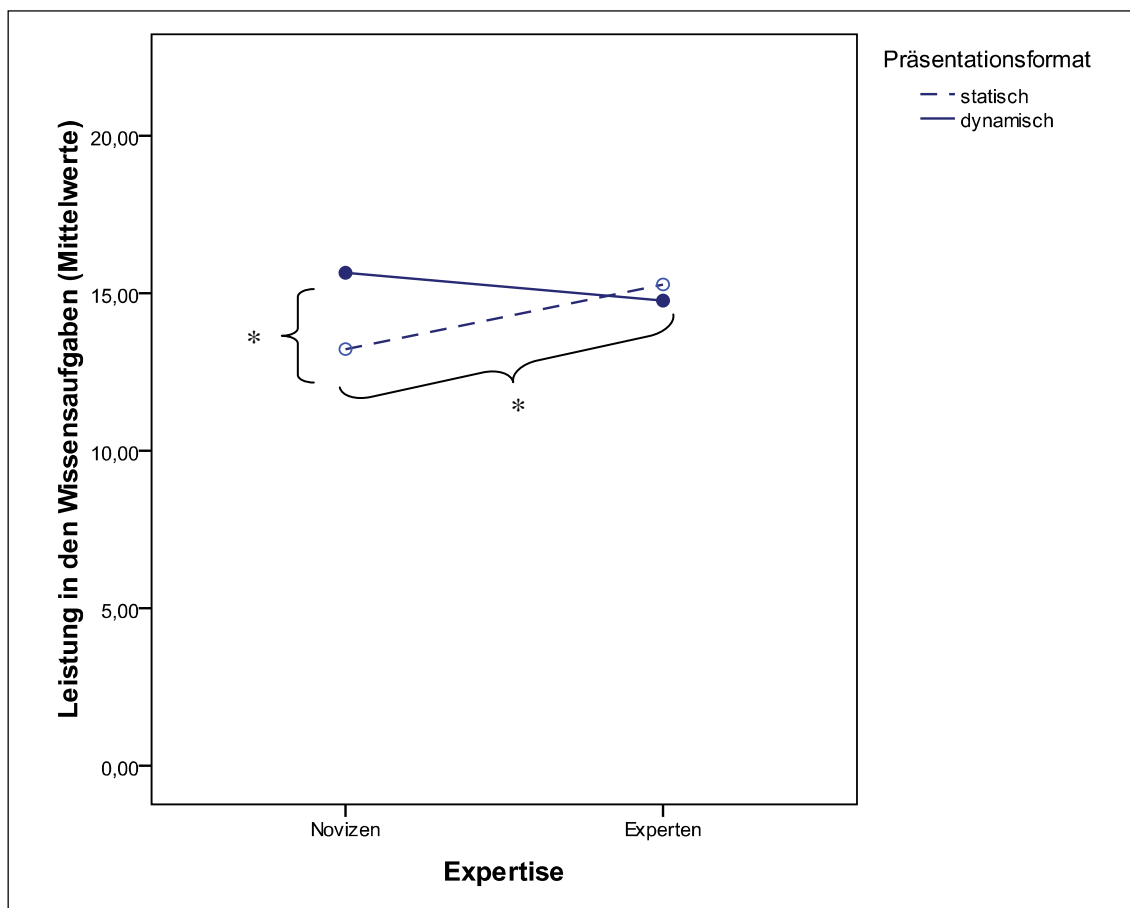


Abbildung 4.8

Mittlere Leistungen von Experten und Novizen in den Wissensaufgaben getrennt für die beiden Präsentationsformen statisch und dynamisch (Maximalpunktzahl: 22, * $p < .05$) (vgl. Malone et al., 2012)

Um die Art der Interaktion zwischen den beiden Faktoren Expertise und Präsentationsform zu spezifizieren, wurden einfache Effekte (*simple effects*) berechnet.

Für die Mehrfachvergleiche wurde eine Anpassung des α -Niveaus nach Bonferroni vorgenommen. Dabei zeigte sich, dass die Novizen in der statischen Version signifikant weniger Aufgaben richtig lösten als in der dynamischen Version ($F(1, 115) = 6.03$; Mittlere Differenz = 2.42; $p = .02$; $\eta_p^2 = .05$), während für die Gruppe der Experten keine signifikanten Mittelwertsunterschiede vorlagen ($F(1, 115) = .30$; Mittlere Differenz = .51; $p = .59$). In der statischen Version des Testmaterials erbrachten die Experten signifikant bessere Leistungen als die Novizen ($F(1, 115) = 4.31$; Mittlere Differenz = 2.05; $p = .04$; $\eta_p^2 = .04$), während sich in der dynamischen Version keine statistisch bedeutsamen Unterschiede zeigten ($F(1, 115) = .85$; Mittlere Differenz = .88; $p = .36$).

Da es sich bei den Wissensaufgaben um eine reliable Skala handelt, ihre Validität aufgrund der nicht sehr deutlichen Mittelwertsunterschiede zwischen den erfahrenen und den unerfahrenen Fahrern jedoch nicht nachgewiesen werden kann, wird möglicherweise ein genaueres Leistungsmaß benötigt. Durch dieses Maß sollten zudem Deckeneffekte, die bei denjenigen Aufgaben mit besonders geringer Schwierigkeit auftauchen können, entgegengewirkt werden. Für das Lernen werden gelegentlich Effizienzmaße berechnet, die die Effizienz beim Lernen wie von Paas und Van Merriënboer (1993) unter Einbezug der Leistung in einem Nachtest und der Höher der mentalen Belastung und/ oder der Trainingszeit (Salden, Paas, Broers & Van Merriënboer, 2004) schätzen. In Anlehnung an diese Effizienzmaße wurde Formel (5) als ein Effizienzmaß für die Bearbeitung von Testaufgaben entwickelt.

$$TE = \frac{P - T}{\sqrt{2}} \quad (5)$$

Die Testeffizienz (TE) ist damit abhängig von der Leistung im Test (P) und der Bearbeitungszeit (T). Hohe Werte resultieren aus hoher Leistung und geringer Bearbeitungszeit und stehen damit für eine hohe Effizienz. Als Leistung im Test wurde die Anzahl der richtig gelösten Wissensaufgaben eingesetzt, die am gemeinsamen Mittelwert von Experten und Novizen über beide Testversionen hinweg z-standardisiert wurden. Für jede Aufgabe wurde zudem die Bearbeitungszeit protokolliert. Diese Zeiten wurden ebenfalls an ihrem gemeinsamen Mittelwert z-standardisiert. Für jede Person ging ihr mittlerer z-Wert in die Formel der Testeffizienz ein. Tabelle 4.8 zeigt die mittleren Effizienzwerte in den verschiedenen Bedingungen.

Tabelle 4.8

Mittelwerte und Standardabweichungen der Testeffizienz der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch)

	Experten/ statische Version n = 32		Novizen/ statische Version n = 27		Experten/ dynamische Version n = 30		Novizen/ dynamische Version n = 31	
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)
Testeffizienz	.59	(.67)	.24	(.98)	-.52	(.58)	-.31	(.79)

Anmerkungen. M = Mittelwert; SD = Standardabweichung

In einer Kovarianzanalyse mit der Kovariate Visualisierungsfähigkeit, den unabhängigen Variablen Expertise und Präsentationsformat und der abhängigen Variable Testeffizienz zeigte sich weder ein signifikanter Haupteffekt der Kovariate ($F(1, 115) = .30; p = .59$) noch der Expertise ($F(1, 115) = .14; p = .71$). Mit der statischen Version des Testmaterials wurde im Durchschnitt effizienter gearbeitet als mit der dynamischen Version des Testmaterials. Es resultierte ein statistisch bedeutsamer Haupteffekt des Präsentationsformates ($F(1, 115) = 35.55; p \leq .001; \eta_p^2 = .24$). Dieser hohe Effekt ist darauf zurück zu führen, dass in der dynamischen Version Videos gezeigt wurden, die selbst schon einige Sekunden der Bearbeitungszeit beanspruchten. Die Videos konnten von den Teilnehmern nicht selbstständig gestoppt werden. Die präsentationsformabhängigen Effizienzunterschiede sind demnach systembedingt. Der Haupteffekt ist zur Beantwortung der Hypothesen nicht von Bedeutung, sondern die Wechselwirkung zwischen Präsentationsformat und Expertise. Tendenziell zeigte sich ein Interaktionseffekt von Expertise und Präsentationsformat ($F(1, 115) = 3.87; p = .05; \eta_p^2 = .03$). Die Effizienzmaße der Novizen scheinen tendenziell in geringerem Maße vom Präsentationsformat abzuhängen als die der Experten.

Anschließend wurde die Leistung in den Einschätzungsaufgaben als abhängige Variable betrachtet. Alle Einschätzungsaufgaben wurden trotz ihrer geringen internen Konsistenz zu einer Skala zusammengefasst. Der Experten-Novizen-Vergleich sollte zeigen, ob die Aufgaben trotz geringer Reliabilität ein gewisses Maß an Kriteriumsvalidität aufweisen. Ein möglicher Einfluss der Faktoren Expertise und Präsentationsform auf die Leistung in den Einschätzungsaufgaben wurde durch eine Kovarianzanalyse mit der Kovariate Visualisierungsfähigkeit ermittelt. Der Einfluss der Kovariate erwies sich als statistisch bedeutsam ($F(1, 115) = 5.07; p = .03; \eta_p^2 = .04$). Es zeigte sich weder ein Haupteffekt der Expertise ($F(1, 115) = 0.05; p = .82$) noch der Präsentationsform ($F(1, 115) = 0.01; p = .98$). Es lag ebenfalls keine Interaktion der Faktoren Expertise und Präsentationsform vor

($F(1, 115) = 0.11$; $p = .74$). Experten lösten demnach nicht mehr Aufgaben als Novizen. Zudem lösten die Personen, denen die statische Version des Testmaterials vorlag etwa gleich viele Aufgaben richtig wie die Personen, die die dynamische Version bearbeiteten.

In den nicht durch Bildmaterial angereicherten Textaufgaben zum Verkehrswissen wurde der Faktor Präsentationsformat nicht variiert, da den Teilnehmern zur Beantwortung dieser Aufgaben kein Bildmaterial vorlag. In diesen Aufgaben waren die Experten ($M = 6.52$; $SD = 2.13$) den Novizen ($M = 5.93$; $SD = 2.28$) in der Anzahl der durchschnittlich richtig gelösten Aufgaben überlegen. Der Mittelwertsunterschied erwies sich in einer einfaktoriellen Kovarianzanalyse mit der Kovariate Visualisierungsfähigkeit jedoch nicht als statistisch signifikant ($F(1, 117) = 0.58$; $p = .45$). Der Einfluss der Kovariate war hingegen statistisch bedeutsam ($F(1, 117) = 11.16$; $p = .001$; $\eta_p^2 = .09$).

4.1.3.5 Einfluss von Unfallvorgeschichte und kritischen Situationen im Verkehr

Zur Untersuchung der Zusammenhänge zwischen der Testleistung in den verschiedenen Aufgabenformaten und dem Unfallrisiko werden ausschließlich die Daten der 63 Experten analysiert. Nur 13 % der Personen gaben an, in den letzten sechs Monaten mindestens einen Unfall verursacht zu haben. Im Durchschnitt über alle Experten hinweg wurden 0.22 Unfälle in diesem Zeitraum verursacht ($SD = .68$). Diejenigen Personen, die in den letzten sechs Monaten Unfälle verursacht hatten, gaben an, es habe sich dabei um ein bis maximal vier leichtere und mittelschwere Unfälle pro Person gehandelt. Unfälle mit Personenschaden wurden nicht berichtet. Aufgrund der geringen Unfallzahlen eignet sich diese Angabe nicht für inferenzstatistische Analysen. Es wurde außerdem erhoben, wie viele Unfälle die Teilnehmer insgesamt seit dem Erwerb ihrer Fahrerlaubnis verursacht hatten. Dabei zeigte sich, dass 30 % der Personen bisher noch keinen und weitere 25 % erst einen Unfall verursacht hatten. Die restlichen Personen gaben an, bisher zwischen zwei und acht Unfälle verursacht zu haben. Dabei handelt es sich lediglich bei 1 % der Unfälle um schwere Unfälle mit Personenschäden. Im Mittel hatten die Teilnehmer bisher 1.73 Unfälle verursacht. Aufgrund der höheren Varianz der Angaben eignet sich diese Variable zur inferenzstatistischen Analyse. Die Anzahl der Unfälle wurde dazu an der Dauer des Führerscheinbesitzes relativiert. Es wurde für jede Versuchsperson die durchschnittliche Unfallzahl pro Jahr ihres Fahrerlaubnisbesitzes berechnet ($M = .24$; $SD = .25$).

Ein weiteres Maß, das Aufschluss über die persönliche Unfallanfälligkeit der Teilnehmer geben soll, ist die Anzahl an kritischen Situationen, die die Teilnehmer in den

letzten vier Wochen im Verkehr erlebt hatten. Die Teilnehmer gaben an, durchschnittlich an 2.56 ($SD = 5.24$) kritischen Verkehrssituationen beteiligt gewesen zu sein.

Zur Bestimmung des Zusammenhangs von Unfallzahlen bzw. kritischen Situationen und der Leistung in den verschiedenen Aufgabenarten wurden Korrelationen berechnet, die in Tabelle 4.9 dargestellt werden.

Lediglich zwischen den Leistungen in den Einschätzungsaufgaben der statischen Version und der Anzahl der kritischen Verkehrssituationen zeigte sich ein signifikanter positiver Zusammenhang.

Tabelle 4.9

Korrelationen (Pearson r) zwischen Unfallzahlen/ Kritischen Situationen und der Leistung in den Wissens-/ Einschätzungs-/ Textaufgaben

	Wissensaufgaben			Einschätzungsaufgaben			Textaufgaben
	statisch	dynamisch	gesamt	statisch	dynamisch	gesamt	gesamt
Unfälle pro Jahr	.18	-.22	-.01	.16	-.11	.05	-.19
Kritische Situationen	.13	.18	.14	.36*	.03	.09	.01

Anmerkungen. $N = 63$; n (statische Version) = 32; n (dynamische Version) = 31. * $p < .05$

4.1.4 Diskussion

In der folgenden Diskussion werden die Ergebnisse zunächst zusammengefasst und hinsichtlich ihrer Bedeutung für die einzelnen aufgestellten Hypothesen bewertet. Anschließend werden Implikationen, die sich aus den Ergebnissen für das weitere empirische Vorgehen ergeben, diskutiert.

4.1.4.1 Bedeutung der Ergebnisse für die Hypothesen

Das Ziel der Studie war es, den Einfluss von Fahrexpertise und Präsentationsformat auf die Leistung in verschiedenen verkehrsspezifischen Aufgabenformaten zu ermitteln. Die verschiedenen Aufgabenformate sollten hinsichtlich ihrer Güte bewertet und damit ihre prinzipielle Eignung zum Einsatz in der theoretischen Fahrerlaubnisprüfung überprüft werden. Analog zu den Befunden der Expertiseforschung wurde vermutet, dass erfahrene Fahrer bessere Leistungen erbringen sollten als Fahrschüler (Hypothese 1). Die Bestätigung dieser Hypothese würde für die Kriteriumsvalidität des betreffenden Aufgabenformates sprechen. Die Ergebnisse der Studie konnten diese Hypothese weder für die

Wissensaufgaben, die aus Aufgaben der theoretischen Fahrerlaubnisprüfung abgeleitet waren, noch für die eigens für diese Studie entwickelten Einschätzungsaufgaben zur Erfassung der Faustregeln stützen.

Experten und Novizen lösten etwa gleich viele Wissensaufgaben mit Bildmaterial und waren in ihrer Aufgabenbearbeitung ähnlich effizient. Auch bei ihren Leistungen in den reinen Textaufgaben zum Verkehrswissen unterschieden sich die beiden Expertisegruppen nicht bedeutsam voneinander. Es liegt nahe, die Ergebnisse dahingehend zu interpretieren, dass das verkehrsbezogene deklarative Wissen nicht generell mit der Fahrerfahrung zu steigen scheint. Da die Fahrexpertise aber mit der Fahrerfahrung steigt, kann auch geschlussfolgert werden, dass die Wissensaufgaben Fahrexpertise nicht abbilden können und damit nicht als kriteriumsvalide bezeichnet werden können. Die Wissensaufgaben entsprachen, abgesehen von dem am Computer erstellten Bild beziehungsweise Video, einigen Aufgaben der derzeitigen theoretischen Fahrerlaubnisprüfung genau. Daher wäre es möglich, dass diejenigen Fahrschüler, die diese Aufgaben bereits explizit geübt hatten, die besten Leistungen erbracht haben. Da dieses explizite Lernen für die Prüfung anhand der Aufgaben bei den Experten schon mindestens zwei Jahre zurückliegen sollte, sollten eher die Fahrschüler die Experten übertreffen. Dass die Gruppe der Novizen dennoch nicht deutlich bessere Leistungen erbracht hat als die Experten, kann auf die Tatsache zurückgeführt werden, dass sich nicht alle Novizen auf dem gleichen Stand ihrer Fahrausbildung befanden und folglich nicht alle die Aufgaben bereits geübt hatten.

Die Bearbeitung der für diese Studie neu erstellten Einschätzungsaufgaben konnte durch die Teilnehmer in Vorfeld nicht trainiert werden. Lediglich die Faustregeln für Anhalteweg, Sicherheitsabstand und Überholweg, deren Kenntnis implizit zur Lösung der Aufgaben beitragen kann, werden im Fahrschulunterricht erworben. Die Ergebnisse bezüglich der Einschätzungsaufgaben stellen deren Sensitivität gegenüber Fahrerfahrung und damit ihre kriteriumsbezogene Validität in Frage, da die Experten den Novizen nicht überlegen waren. Dies liegt vermutlich teilweise in der Tatsache begründet, dass die Aufgaben nach den Faustformeln zu Anhalte-, Überholweg und Sicherheitsabstand erstellt wurden, wie sie als Regeln in der Fahrschule gelehrt werden. Diese Regeln entsprechen jedoch vermutlich nicht den realen Erfahrungen vieler langjähriger Fahrer. So legt man beispielsweise mit einem neuen Auto bei einer Vollbremsung einen deutlich geringeren Weg zurück als durch die Berechnung des Anhaltewegs vorhergesagt wird. Aufgrund dessen wurden die Regeln, die in der Fahrschule erlernt wurden, vermutlich durch die Erfahrung nicht internalisiert.

Wahrscheinlicher ist die Möglichkeit, dass Erfahrungswerte sie vollständig ersetzt haben. In der Studie wurde nicht danach gefragt, ob die der Berechnung zu Grunde gelegte Regel von den Probanden verwendet wurden. Ein weiterer Kritikpunkt am Aufgabenmaterial selbst ist, dass es dem geübten Fahrer bei den Aufgaben zum Überholweg (Block 3) realitätsfern erscheinen sollte, seine Geschwindigkeit beim Überholen konstant zu halten. In der Realität beschleunigt man beim Überholen kontinuierlich, um diesen risikobehafteten Vorgang möglichst kurz zu gestalten. Diese Diskrepanz zur Realität wurde auch von einigen Versuchsteilnehmern rückgemeldet. Dieser Umstand lässt Zweifel an der inhaltlichen Validität der Aufgaben aufkommen.

Neben dem Faktor Expertise wurde in dieser Studie der Faktor Präsentationsformat experimentell variiert. Dabei wurde vermutet, dass die Novizen bei der Aufgabenbearbeitung von einer dynamischen Darstellung des Bildmaterials im Gegensatz zu den Experten profitieren würden (Hypothese 2). Es konnte gezeigt werden, dass die Leistung in den Wissensaufgaben insgesamt unabhängig davon zu sein scheint, ob die Aufgaben mit einem Bild oder einem Video zur Veranschaulichung einer Verkehrssituation ausgestattet waren oder nicht. Es konnte aber eine Wechselwirkung der Faktoren Expertise und Präsentationsform festgestellt werden. Durch die Betrachtung der einfachen Effekte kann für die Wissensaufgaben Hypothese 2 gestützt werden, denn für die Gruppe der Novizen zeigte sich, dass in der dynamischen Version des Testmaterials bessere Leistungen erbracht wurden als in der Version mit statischem Bildmaterial. Für die Experten hingegen konnte kein von der Präsentationsform abhängiger Leistungsunterschied festgestellt werden. Die Tatsache, dass die Teilnehmer ohne Fahrerfahrung von einer dynamischen Darstellung der Verkehrsszenarien profitierten, kann dadurch erklärt werden, dass Novizen sich den Verlauf der Situation anhand eines statischen Bildes nicht so gut vorstellen können, als wenn ihnen die mentale Animation des Szenarios durch eine externe dynamische Darstellung abgenommen wird. Im Falle der Experten scheint es hingegen keinen Unterschied zu machen, ob die Aufgaben dynamische Präsentationen der Verkehrsszenarien beinhalten oder nicht. Das Ergebnis kann als *Effekt des Vorwissens (prior knowledge principle, Kalyuga, 2005)* interpretiert werden, der besagt, dass sich eine gute Gestaltung des Test- beziehungsweise Lernmaterials nur bei geringer Expertise positiv auf die Leistung auswirkt. Begründet ist diese Annahme auf der Tatsache, dass die Ressourcen des Arbeitsgedächtnisses als limitiert gelten. Aufgrund ihrer Erfahrung sind Experten in der Lage, aus einem Bild den Verlauf einer Situation zu erschließen. Die Computeranimation stellt im besten Fall redundante

Informationen zur Verfügung oder könnte sogar dazu führen, dass Experten eine Diskrepanz zur Realität bemerken. Beides kann die Ressourcen des Arbeitsgedächtnisses belasten und somit die Aufgabenbearbeitung beeinträchtigen, was in der Folge zu einem so genannten *expertise-reversal-Effekt* führen kann. Dabei erbringen Experten sogar schlechtere Leistungen, wenn das Lernmaterial den Gestaltungsprinzipien zufolge gut gestaltet ist, als wenn es sich um weniger gut gestaltetes Lernmaterial handelt (vgl. z.B. Kalyuga et al., 2003). Dass die Experten in dieser Studie nicht von der dynamischen Darstellung der Verkehrsszenarien profitiert haben, könnte also dadurch begründet sein, dass sie Redundanzen und Diskrepanzen zur Wirklichkeit, die sie wahrnehmen, verarbeiten mussten, was die eigentliche Aufgabenbearbeitung gestört haben könnte. Das Ergebnis der Studie, dass Novizen von der dynamischen Darstellung der Verkehrssituation profitieren, weist darauf hin, dass das dynamische Präsentationsformat die Instruktion, das Aufgabenverständnis und damit die Bearbeitung der Aufgaben für die Novizen erleichtert. Dies spricht dafür, dass die Reliabilität der Aufgaben durch den Einsatz von Dynamik gesteigert werden kann und damit die zuverlässige Messung von Fahrkompetenz begünstigt wird. Dagegen ist einzuwenden, dass sich bei den paarweisen Vergleichen außerdem zeigte, dass die Experten die Novizen in der statischen Version des Testmaterials bedeutsam übertrafen. In ihrer statischen Version scheinen die Aufgaben demnach kriteriumsvalid zu sein, während sie es in ihrer dynamischen Version nicht sind.

Ob die Aufgaben durch die dynamische Darstellung einer Verkehrssituation angereichert sind oder nicht, schien sich auch im Falle der Einschätzungsaufgaben nicht generell bedeutsam auf die Leistung der Teilnehmer auszuwirken. Es zeigte sich für die Einschätzungsaufgaben keine Interaktion zwischen den Faktoren Expertise und Präsentationsformat. Für diesen Aufgabentyp konnte demnach die Annahme, dass die Novizen von einer dynamischen Darstellung des Bildmaterials profitieren, durch die Ergebnisse nicht gestützt werden.

Als ein weiteres Kriterium für die Validität der Aufgabenformate wurde ein negativer Zusammenhang zwischen Unfallanfälligkeit und Leistung im jeweiligen Aufgabenformat angenommen (Hypothese 3). Diese Hypothese konnte durch die Ergebnisse weder für die Wissens- noch für die Einschätzungsaufgaben gestützt werden, denn es zeigten sich keinerlei bedeutsame negative Korrelationen der retrospektiv berichteten Unfallhäufigkeit der Teilnehmer oder der Häufigkeit der von ihnen verursachten kritischen Verkehrssituationen mit ihrer Leistung in den einzelnen Aufgabenformaten. Der Befund, dass Personen mit mehr

selbstverschuldeten Unfällen in der Vergangenheit nicht generell schlechter in den Wissensaufgaben abschnitten, legt nahe, dass das Risiko zu verunfallen, nicht mit deklarativem Verkehrswissen zusammenhängt. Unfallstatistiken zeigen, dass unangepasste Geschwindigkeit, zu niedriger Sicherheitsabstand und unangemessenes Überholen häufig zu Unfällen führen (Statistisches Bundesamt, 2008, 2010b) und demnach die richtige Einschätzung von Verkehrssituationen dahingehend besonders sicherheitsrelevant sein sollte. Dass die Einschätzungsaufgaben dies nicht abbilden konnten, ist vermutlich in ihrer zuvor beschriebenen mangelnden inhaltlichen Validität begründet.

Zusätzlich zu den Hypothesen wurde die Forschungsfrage formuliert, inwieweit die Visualisierungsfähigkeit mit den Leistungen in den unterschiedlichen Aufgabenformaten zusammenhängt. Es zeigte sich, dass die Experten die Novizen im Paper Folding Test deutlich übertrafen. Daher wurde die Visualisierungsfähigkeit als Kovariate in die durchgeführten Varianzanalysen einbezogen. Es zeigte sich jedoch weder ein Haupteffekt der Fähigkeit zur räumlichen Visualisierung auf die Leistung in den Wissensaufgaben mit Bildmaterial noch auf die Leistung in den Einschätzungsaufgaben. Allerdings konnte ein schwacher statistisch bedeutsamer korrelativer Zusammenhang zwischen der Visualisierungsfähigkeit und der Leistung in den Einschätzungsaufgaben festgestellt werden. Ein deutlicherer Zusammenhang mit der Visualisierungsfähigkeit und der Leistung konnte für die reinen Textaufgaben ermittelt werden. Dieser Befund stützt die ability as compensator-Hypothese der Instruktionspsychologie, die besagt, dass Personen mit stark ausgeprägten Fähigkeiten (hier Visualisierungsfähigkeit) die Schwächen des Präsentationsmaterials (hier fehlende Bild- oder Videoinformationen) ausgleichen können (Mayer & Sims, 1994). Eine Erklärung für die Ergebnisse besteht demnach darin, dass Personen mit einer hohen Visualisierungsfähigkeit in der Lage sind, sich die in Textform beschriebene Verkehrssituation ohne Bild- oder Videounterstützung adäquat vorzustellen, während Personen mit geringeren Visualisierungsfähigkeiten dies weniger gut können. Dafür spricht zudem die Tatsache, dass bei den Aufgaben mit visueller Unterstützung durch Bildmaterial kein hoher Zusammenhang zwischen der Leistung in den verkehrsspezifischen Aufgaben und der Visualisierungsfähigkeit auftrat.

4.1.4.2 Implikationen für das weitere empirische Vorgehen

Durch die dargestellte erste Vorstudie sollten valide und reliable Aufgabenformate zur Messung von Aspekten der Verkehrskompetenz identifiziert werden, die in eine spätere Längsschnittstudie eingehen sollten. Es sollte zudem entschieden werden, welches

Präsentationsformat (statisch oder dynamisch) für die ausgewählten Aufgaben angemessen erscheint.

Die in der Studie eingesetzten Wissensaufgaben mit Bildmaterial ermöglichten keine Trennung von Experten und Novizen, was gegen ihre Kriteriumsvalidität spricht. Dies wurde allerdings auf die Tatsache zurückgeführt, dass die meisten Novizen diese Aufgaben für ihre theoretische Prüfung geübt haben könnten und nur daher nicht weniger Aufgaben richtig gelöst haben als die Experten. Es ist daher zu erwarten, dass die Wissensaufgaben besser zwischen den Experten und den Novizen zu trennen vermögen, wenn es sich bei den Novizen um Personen handelt, die sich noch ganz am Beginn ihrer Fahrschul Ausbildung befinden, was zum ersten Messzeitpunkt der geplanten Längsschnittstudie der Fall sein sollte. Da die Aufgaben außerdem hinsichtlich ihrer Reliabilität überzeugen konnten, werden die Wissensaufgaben mit Bildmaterial für die Längsschnittuntersuchung ausgewählt. Die Entscheidung, ob sie dabei in ihrer statischen oder dynamischen Version dargeboten werden sollen, muss differenziert betrachtet werden. Durch die dynamische Darstellung der Verkehrsszenarien konnte die Reliabilität der Aufgaben gesteigert werden, da die Dynamik die Novizen in ihrem Verständnis der Situation unterstützte, was den Messfehler verringern sollte. Dabei handelt es sich zudem um einen Validitätsaspekt, da durch das gesteigerte Verständnis der Aufgabe ein Testbias verhindert werden kann. Die Aufgabebearbeitung wird also nicht durch den Umstand gestört, dass die Teilnehmer sich die betreffende Verkehrssituation nicht richtig vorstellen können. Außerdem wird die inhaltliche Validität der Aufgaben durch die realistischere Darstellungsweise gefördert. In der statischen Version des Testmaterials hingegen war aufgrund der Überlegenheit der Experten gegenüber den Novizen Kriteriumsvalidität gegeben. Für die Entscheidung darüber, in welchem Präsentationsformat die Aufgaben in der weiterführenden Studie dargeboten werden sollten, ergibt sich hieraus ein Konflikt. Durch die beiden Präsentationsformate werden oberflächlich betrachtet unterschiedliche Güteaspekte gesichert. Es hat sich gezeigt, dass die Novizen von einer dynamischen Darstellung der Verkehrsszenarien profitieren, was darauf hindeutet, dass sie die Situation bei bewegter Darbietung besser verstehen konnten. Das bedeutet im Umkehrschluss, dass die Novizen in den Aufgaben, die in der statischen Version dargeboten wurden, möglicherweise schlechtere Leistungen erbrachten, weil ihr Verständnis für die Situation geringer war. Aus dieser Überlegung folgt, dass der Unterschied zwischen den Experten und den Novizen in der statischen Version nicht aufgrund der Aufgabenvalidität in dieser Version deutlicher ausfiel, sondern durch das unterschiedlich hohe Instruktionsverständnis der beiden

Expertisegruppen. Unterstützt man das Instruktionsverständnis der Novizen durch Dynamik, wird sichergestellt, dass ihre Leistung in den Aufgaben weniger davon abhängt, ob sie sich die Situation vorstellen können, als davon inwieweit sie über verkehrsspezifisches Wissen verfügen, um die jeweilige Aufgabe zu lösen. Die Steigerung des Instruktionsverständnisses durch das dynamische Präsentationsformat wirkt sich demnach positiv auf die Reliabilität und Validität der Aufgaben aus. Die Aufgaben werden daher in der Längsschnittstudie in ihrer dynamischen Präsentationsform beibehalten, da durch diese Darstellungsform die Güte der Aufgaben deutlich gefördert wurde.

Die Einschätzungsaufgaben haben sich empirisch weder als valide noch als reliabel gezeigt und werden daher nicht in die Längsschnittstudie einbezogen. Ebenso wird in der Längsschnittstudie auf den Einsatz der reinen Textaufgaben verzichtet, da sie sich nicht als kriteriumsvalide erwiesen hatten und zudem der relativ deutliche Zusammenhang mit der Visualisierungsfähigkeit für Einschränkungen hinsichtlich der Güte der Aufgaben spricht.

4.2 Vorstudie 2

4.2.1 Ziele und Fragestellungen

Die Leistung in den in Vorstudie 1 eingesetzten Aufgabenformaten war nicht fahrexpertiseabhängig, was gegen ihre kriteriums-basierte Validität spricht. In der zweiten Vorstudie wurde ebenfalls das Ziel verfolgt, durch einen Experten-Novizen-Vergleich ein computerbasiertes Aufgabenformat hinsichtlich seiner Kriteriumsvalidität zu testen. Zudem sollte auch in Vorstudie 2 ein Vergleich von Aufgaben mit statischem und dynamischem Bildmaterial erfolgen, um zu ermitteln, welches Präsentationsformat die Trennung von Experten und Novizen begünstigt.

Zur Auswahl geeigneter Aufgaben- und Antwortformate wurden die Befunde der Expertise- und Verkehrssicherheitsforschung herangezogen. In der Expertiseforschung konnte gezeigt werden, dass Experten den Novizen umso deutlicher überlegen sind, je spezifischer die gestellte Aufgabe für die jeweilige Domäne ist (Glaser & Chi, 1988). Die in Vorstudie 1 eingesetzten Aufgaben erfassten verkehrsspezifisches Faktenwissen mithilfe von Mehrfachwahlaufgaben. Die abgeprüften Inhalte waren demnach spezifisch für die Domäne Autofahren. Die Verhaltensantwort, die durch das Mehrfachwahlantwortformat von den Teilnehmern gefordert wurde, steht allerdings in keinem Zusammenhang mit dem tatsächlichen Verhalten beim Fahren und war daher nicht spezifisch für die Domäne Autofahren. Implizite Reaktionen, die beim Fahren eine große Rolle spielen, können durch

eine retrospektive Befragung, wie sie z.B. bei Mehrfachwahlaufgaben nach der Präsentation des zugehörigen Bildes oder Videos erfolgt, nicht abgeprüft werden (Rauch et al., 2009). Besteht das Ziel darin, Aufgaben zu erstellen, die kriteriumsvalid zwischen Fahrexperten und -novizen trennen, wird die Entwicklung eines handlungsnäheren und damit domänenspezifischeren Antwortformates, das auch implizites, intuitives Verhalten erfasst, erforderlich. Dazu wurde in der zweiten Vorstudie ein Aufgabenformat mit Reaktionszeitmessung, ähnlich einem *Hazard-Perception-Test* (vgl. Kapitel 2.3) entwickelt. Das dabei geforderte schnelle Reagieren auf Verkehrsgegebenheiten sollte eher dem Verhalten eines Fahrers im Realverkehr entsprechen als klassische gebundene und freie Antwortformate und damit spezifischer für die Domäne Autofahren sein.

Bisher sind einige Studien bekannt, in denen das Reaktionszeitaufgabenformat zur Messung der Gefahrenwahrnehmung eingesetzt wurde und sich dabei als kriteriumsvalid erwiesen hat (vgl. Kapitel 2.3.3). Dabei wurden entweder Fotos oder Videos mit Gefahrenhinweisen verwendet, die die Teilnehmer so schnell wie möglich erkennen mussten (vgl. Kapitel 2.3.4). In diesen Studien wurde somit *entweder* mit statischem Bildmaterial *oder* dynamischem Bildmaterial gearbeitet. Keine der Studien bezog dabei exakt die gleichen Verkehrsszenarien in einer statischen und in einer dynamischen Version ein, um sie miteinander zu vergleichen. Dies wurde in der vorgestellten Studie geleistet, um zu bestimmen, welches Präsentationsformat zur Trennung von Experten und Novizen geeigneter ist.

Analog zu Vorstudie 1 werden als Experten diejenigen Teilnehmer bezeichnet, die seit mindestens zwei Jahren eine Fahrerlaubnis der Klasse B besitzen. Personen ohne Fahrerlaubnis der Klasse B werden als Novizen bezeichnet. Im Folgenden werden auf der Grundlage der Befunde aus Expertiseforschung und der Forschung zur Gefahrenwahrnehmung im Verkehr Hypothesen für die Vorstudie 2 abgeleitet.

Einen über verschiedene Domänen stabilen Befund der Expertiseforschung stellt das Ergebnis dar, dass Experten innerhalb „ihrer“ Domäne bessere Leistungen als Novizen erbringen (Glaser & Chi, 1988). Experten zeichnen sich unter anderem dadurch aus, dass sie spezifische Probleme zuverlässiger und schneller lösen als Novizen. Cellier, Eyrolle und Marine (1997) stellen in ihrer Überblicksarbeit heraus, dass Experten im Vergleich zu Novizen Probleme eher antizipieren und somit präventiv statt reaktiv auf Störungen und Schwierigkeiten reagieren können. Expertise wird den meisten Theorien zum Expertiseerwerb zufolge durch praktische Übung aufgebaut (vgl. Kapitel 2.1.3). Das Gleiche gilt offensichtlich

für die Domäne Autofahren: Erfahrene Fahrer fahren besser und sicherer als unerfahrene Fahrer, was sich in der relativ geringen Unfallrate der erfahrenen Fahrer widerspiegelt (vgl. Kapitel 2.2). Fahrer mit Fahrpraxis sollten demnach, aufgrund ihrer Erfahrung, Fahrschüler in Aufgaben, die Aspekte von Fahrexpertise messen, deutlich übertreffen.

Das Reaktionszeitaufgabenformat scheint Aspekte von Fahrexpertise zu erfassen, da die Leistung in derartigen Aufgaben im Zusammenhang mit Fahrpraxis und Unfallbeteiligung steht (vgl. Kapitel 2.3.3). Fahrexperten können gefährliche Situationen im Verkehr zuverlässig antizipieren und sind den unerfahrenen Fahrnovizen daher in *Hazard-Perception*-Aufgaben, bei denen das Absuchen der Verkehrsumwelt nach Gefahrenreizen verlangt wird, überlegen.

Hypothese 1: Experten lösen mehr Aufgaben richtig als Novizen.

Hypothese 2: Experten reagieren schneller auf Bremshinweise als Novizen.

Ein Vorteil von computerbasiertem Training und computerbasierter Testung besteht in der Möglichkeit Lern- und Testinhalte, die sich in der Realität bewegen, auch bewegt darzustellen. Die Forschung zum Lernen mit Neuen Medien konnte zeigen, dass dynamische Präsentationen unter bestimmten Umständen – nämlich besonders wenn prozedurales statt deklaratives Wissen erworben werden soll – das Lernen begünstigen können (Höffler & Leutner, 2007). Inwiefern diese Befunde auch auf das Lösen von Reaktionszeitaufgaben im Verkehrskontext übertragbar sind, wurde bisher noch nicht hinreichend untersucht. Bei der Erforschung der Gefahrenwahrnehmung wurden zwar Studien durchgeführt, die die Reaktion von Probanden auf Reize innerhalb entweder unbewegter oder bewegter Darstellungen von Verkehrsszenarien erfassten, jedoch Studien, die das Präsentationsformat systematisch verglichen haben, sind bisher nicht bekannt. In Vorstudie 1 konnte gezeigt werden, dass das Präsentationsformat zwar keinen generellen Einfluss auf die Testleistung hatte, sich jedoch im Falle der eingesetzten Wissensaufgaben in Abhängigkeit vom Faktor Expertise auswirkte. Novizen schienen von einer dynamischen Darstellung der Verkehrsszenarien, auf die sich die gestellten Fragen bezogen, zu profitieren. Dieses Ergebnis kann allerdings nicht unmittelbar auf die Reaktionszeitaufgaben übertragen werden. Hierbei dient die bildhafte Darstellung nicht lediglich zur Verdeutlichung der Verkehrssituation, sondern aus ihr selbst ergeben sich die Anforderungen an die Teilnehmer. Der Einsatz von dynamischen anstatt statischen Darstellungen von Verkehrsszenarien sollte die Aufgaben realistischer machen. Zum einen sollten die Testteilnehmer die gezeigten bewegten Bilder in ähnlicher Weise wahrnehmen,

wie die ebenfalls durch Bewegung charakterisierten visuellen Eindrücke beim realen Fahren. Zum anderen sollten aber auch die Anforderungen, die das dynamische Präsentationsformat an die Verhaltensantwort der Teilnehmer stellt, den Anforderungen an die Fahrer beim Fahren im Realverkehr eher entsprechen als es bei den Aufgaben im statischen Präsentationsformat der Fall ist. Von den Teilnehmern wird in der dynamischen Version verlangt, innerhalb des richtigen Zeitfensters zu reagieren, während sich ihre Umwelt pausenlos verändert, wie es auch beim realen Fahren der Fall ist. Bei einem statischen Bild ändert sich die Darstellung des Szenarios nicht während der Betrachtung und stellt somit nur einen Ausschnitt aus einer Verkehrssituation dar, die über einen längeren Zeitraum in dieser Konstellation betrachtet werden kann, was beim realen Fahren nicht möglich ist. Daher kann angenommen werden, dass diejenigen Aufgaben, die dynamische Darstellungen von Verkehrsszenarien beinhalten, für die Domäne Autofahren spezifischer sind als Aufgaben, die statische Abbildungen beinhalten. Ein Befund der Expertiseforschung, die so genannte Domänenspezifität, besagt, dass Experten Novizen umso deutlicher übertreffen, je spezifischer die bearbeiteten Aufgaben für die betreffende Domäne sind (Glaser & Chi, 1988; Opwis et al., 1990). Daher kann man annehmen, dass der Einsatz dynamischer Darstellungen in Reaktionszeitaufgaben diese valider macht, was eine Trennung zwischen Experten und Novizen begünstigen sollte.

Hypothese 3: Experten übertreffen Novizen in der Anzahl der identifizierten Bremshinweise deutlicher in der dynamischen als in der statischen Version.

Hypothese 4: Experten übertreffen Novizen in der Geschwindigkeit bei der Identifikation von Bremshinweisen deutlicher in der dynamischen als in der statischen Version.

Die klassischen Verkehrssituationen, die in *Hazard-Perception*-Tests eingesetzt werden, stellen relativ saliente Gefahrensituationen dar, die so schnell wie möglich erkannt werden sollen. In diesen Szenarien tauchen explizite Gefahrenhinweise auf. Zur Lösung dieser Aufgaben ist verkehrsspezifisches Regelwissen für gewöhnlich nicht erforderlich. Biermann (2007) setzte in ihrer Studie zur Gefahrenwahrnehmung Fotos ein, auf denen Gefahrenhinweise von unterschiedlicher Salienz abgebildet waren. Sie konnte zeigen, dass mit steigender Deutlichkeit der Hinweise (abgestuft in drei Kategorien), die Anzahl der durch die Teilnehmer gezeigten angepassten Bremsreaktionen anstieg. Es kann daher vermutet werden, dass Szenarien, die eher explizite statt implizite Hinweise darauf enthalten, dass der Fahrer seine Geschwindigkeit reduzieren sollte, weniger komplex sind und daher eher richtig und schneller darauf reagiert werden kann. Für Vorstudie 2 wurden neben Szenarien mit

expliziten Gefahrenhinweisen auch solche Aufgaben erstellt, zu deren Lösung der Einsatz von Verkehrsregeln – wie z.B. bei der Regel „rechts vor links“ – unbedingt notwendig sind. Diese Verkehrsregeln können eher implizite Bremsgründe darstellen, die in einer komplexen Verkehrssituation nicht sofort ersichtlich sind. Diejenigen Aufgaben, die die Kenntnis von Verkehrsregeln und Verhaltensvorschriften zur Lösung voraussetzen, sollten insgesamt schwieriger sein als Aufgaben, die als kritischen Reiz eine saliente Gefahr beinhalten, auf die reagiert werden muss.

Hypothese 5: Es werden weniger Aufgaben richtig gelöst, zu deren Lösung der Einsatz von Verkehrsregeln notwendig ist, als Aufgaben, zu deren Lösung der Einsatz von Verkehrsregeln nicht notwendig ist.

Hypothese 6: Innerhalb von Aufgaben, die einen salienten Gefahrenhinweis enthalten wird schneller auf die kritische Situation reagiert als innerhalb von Aufgaben, zu deren Lösung die Kenntnis von Verkehrsregeln erforderlich ist.

Der kognitiven Flexibilisierungstheorie von Spiro et al. (1988, 1991) zufolge, wird Expertise durch den Einsatz des erlernten Wissens und der erlernten Fähigkeiten in unterschiedlichen Kontexten und aus unterschiedlichen Perspektiven erworben. Das domänenspezifische Wissen ist bei Experten zum einen umfangreicher aber zum anderen auch besser vernetzt als bei Novizen. Sie können es daher flexibel einsetzen und schneller abrufen. Durch ihre in der Regel reichhaltigen Erfahrungen sollten Fahrexperten die Möglichkeiten gehabt haben, ihr verkehrsspezifisches Wissen oft und in unterschiedlichen Kontexten einzusetzen. Daraus kann gefolgert werden, dass Experten komplexe Verkehrssituationen besser und schneller überblicken und die richtigen Schlüsse für ihre eigene Handlungsplanung in einer Situation ziehen können als Novizen, denen es an reichhaltigen Erfahrungen im Verkehrskontext noch mangelt. Aufgaben, die den Einsatz von Verkehrswissen erfordern (implizite Gefahrenhinweise) sollten spezifischer für den Verkehrskontext sein als Aufgaben, die eine solche Wissensbasis nicht erfordern (explizite Gefahrenhinweise). Den Befunden der Expertiseforschung bezüglich der Domänenspezifität zufolge, sollten Experten den Novizen besonders in solchen Situationen, die ein schnelles Überblicken der komplexen Verkehrssituation und die Ableitung von situationsspezifischen Verhaltensvorschriften verlangen, überlegen sein.

Hypothese 7: Experten übertreffen Novizen bei der Identifikation von Bremshinweisen stärker, wenn zur Lösung der Aufgabe Verkehrsvorschriften einzubeziehen sind.

Hypothese 8: Experten übertreffen Novizen in der Geschwindigkeit der Identifikation von Bremshinweisen stärker, wenn zur Lösung der Aufgabe Verkehrsvorschriften einzubeziehen sind.

Jonah (1996) stellte in einem Überblicksartikel Studien zusammen, in denen Zusammenhänge zwischen verschiedenen Aspekten risikoreichen Fahrverhaltens und dem Persönlichkeitsmerkmal *Sensation Seeking* untersucht wurden. In den meisten Studien zeigten sich positive Zusammenhänge zwischen risikoreichem Fahrverhalten und der Ausprägung des Merkmals *Sensation Seeking*. Die deutlichsten Zusammenhänge zeigten sich für die Unterskala *Thrill and Adventure Seeking (TAS)*. Wie in Kapitel 2.3.1 dargestellt, spielen bei der Gefahrenwahrnehmung auch Bewertungsprozesse eine Rolle, die möglicherweise in einem Zusammenhang mit Persönlichkeitsmerkmalen stehen. Eine Überprüfung des Zusammenhangs zwischen dem Merkmal *Sensation Seeking* und der Fähigkeit zur Gefahrenwahrnehmung steht noch aus.

*Forschungsfrage: Besteht ein Zusammenhang zwischen der Ausprägung des Persönlichkeitsmerkmal *Sensation Seeking* und der Fähigkeit zur Identifikation von Bremshinweisen?*

4.2.2 Methode

4.2.2.1 Experimentelles Design und Stichprobe

Der Studie lag ein 2x2x3-Design mit den Faktoren Expertise (Experten vs. Novizen), Präsentationsform (statisch vs. dynamisch) und dem Messwiederholungsfaktor Aufgabentyp (keine Bremshinweise x explizite Bremshinweise x implizite Bremshinweise) zugrunde. Als abhängige Variablen wurden zum einen die Anzahl richtig gelöster Aufgaben und zum anderen die Reaktionszeit, nach der Präsentation eines Bildes mit Hinweisreiz bzw. nach dem Auftauchen eines Hinweisreizes innerhalb eines Videos, erfasst.

Beim Lernen mit statischem vs. dynamischem Bildmaterial zeigen sich im Durchschnitt mittlere Effekte des Präsentationsformates (Höffler & Leutner, 2007). Die Effektstärken des Faktors Expertise variieren zwischen verschiedenen Studien zur Gefahrenwahrnehmung deutlich. Dies kann zum einen durch die uneinheitliche

Operationalisierung der unabhängigen Variable Expertise begründet werden. Zum anderen unterscheidet sich auch die Gestaltung des Testmaterials zwischen den verschiedenen Studien mitunter deutlich. Daher wird zur Berechnung der optimalen Stichprobe eine mittlere Effektstärke angenommen. Die optimale Stichprobengröße wurde nach Faul, Erdfelder, Lang und Buchner (2007) berechnet. Dabei wurde das Signifikanzniveau auf 5 % festgelegt und es sollte ein mittlerer Effekt mit einer Wahrscheinlichkeit von 90 % gefunden werden. Da es sich bei einem Faktor des Designs um einen dreistufigen Between-Faktor handelt, musste zudem festgelegt werden, inwiefern Zusammenhänge zwischen den Werten der einzelnen Faktorstufen vorausgesagt werden. Aus diesen Überlegungen resultiert für diese Studie eine optimale Gesamtstichprobengröße von $N = 100$. Für Berechnungen der Haupt- und Interaktionseffekte mithilfe einer Varianzanalyse ist es notwendig, eine Gleichbesetzung der vier Zellen zu gewährleisten (Bortz, 2005). Um Ausfällen, die durch die erstmalige Anwendung der Reaktionszeitmessung entstehen konnten, vorzubeugen, wurden etwas mehr Versuchspersonen rekrutiert. An Vorstudie 2 nahmen insgesamt 140 Versuchspersonen teil. In Tabelle 4.10 wird die Verteilung der Novizen und Experten auf die einzelnen Bedingungen dargestellt.

Tabelle 4.10
Verteilung der Versuchsteilnehmer auf die Bedingungen (Vorstudie 2)

Vorstudie 2 N = 140	Präsentationsform	
	statisch	dynamisch
Novizen n = 64	n = 32 (44%)	n = 32 (34%)
Expertise		
Experten n = 76	n = 33 (24%)	n = 43 (21%)

Anmerkung. Die Werte innerhalb der Klammern repräsentieren den Anteil männlicher Probanden getrennt für den Faktor Expertise in der jeweiligen Bedingung.

Insgesamt konnten vier Novizen (zwei statische Version, zwei dynamische Version) aufgrund eines technischen Fehlers die Reaktionszeitaufgaben nicht bearbeiten. Ihre Daten wurden aus der Analyse ausgeschlossen.

Das Lebensalter der Teilnehmer variierte innerhalb der gesamten Stichprobe zwischen 15 und 58 Jahren ($M = 22.32$; $SD = 6.15$). Eine Person war erst 15 Jahre alt, nahm jedoch am Fahrschulunterricht für PKW teil, um zunächst eine Fahrerlaubnis für Kleinkrafträder zu erwerben. Das durchschnittliche Alter der Novizen war 21.83 Jahre ($SD = 8.23$). Die Experten

waren im Durchschnitt 26.92 Jahre alt ($SD = 3.80$) und damit nicht bedeutsam älter als die Novizen ($T(134) = 1.68$; $p = .10$). Das Durchschnittsalter der Novizen unterschied sich nicht bedeutsam zwischen beiden Testbedingungen ($T(56) = 1.29$; $p = .20$). Für die Experten konnten ebenfalls keine bedeutsamen Altersunterschiede zwischen den beiden Testbedingungen festgestellt werden ($T(76) = .52$; $p = .60$). Ein Großteil der Versuchsteilnehmer (65%) studierte zum Zeitpunkt der Untersuchung an der Universität des Saarlandes. Die meisten studentischen Teilnehmer studierten das Fach Psychologie (40.7 %).

Insgesamt hatten 61 Personen zum Testzeitpunkt noch nicht die Fahrerlaubnis für Klasse B erworben und waren somit als Novizen zu bezeichnen. Die Novizen waren allesamt Fahrschüler und hatten in unterschiedlichem Umfang den Theorie- und Praxisunterricht der Fahrschule besucht. 76 Teilnehmer werden als Experten bezeichnet, da sie zum Zeitpunkt der Studie seit mehr als zwei Jahren im Besitz einer Fahrerlaubnis für Klasse B waren. Um zu gewährleisten, dass die Experten seit dem Erwerb ihrer Fahrerlaubnis ein ausreichendes Maß an Fahrpraxis erworben hatten, wurden zum Experiment nur Personen eingeladen, die bis zur Erhebung bereits mindestens 5000 Kilometer gefahren waren. Die Experten gaben an, durchschnittlich maximal 600 Kilometer in der Woche und maximal 50 000 Kilometer im Jahr mit dem PKW zurückzulegen. Insgesamt 22 Personen berichteten bereits mindestens einen Unfall verursacht zu haben, davon hatten zwei Teilnehmer bereits einen Unfall ausgelöst, bei dem es zu einem Personenschaden gekommen war. Die restlichen 20 Personen berichteten von leichteren Unfällen ohne Personenschaden, die sie verursacht hatten.

4.2.2.2 Versuchsmaterial

Innerhalb dieser Studie bearbeiteten die Teilnehmer computerbasierte verkehrsspezifische Reaktionszeitaufgaben (vgl. Anhang B). Zudem wurden den Teilnehmern ein papierbasierter Fragebogen zu demographischen Daten, zu verkehrsbezogenen Angaben sowie ein Fragebogen zur Sensationssuche vorgelegt. Die verschiedenen Testteile werden im Folgenden beschrieben.

Reaktionszeitaufgaben

Es wurden insgesamt 32 Aufgaben erstellt, bei denen eine Messung der Reaktionszeit der Probanden auf relevante Verkehrsgegebenheiten auf Bildern bzw. in Animationen erfolgen sollte. Bei einer der Aufgaben, die immer als erste präsentiert wurde, handelte es sich um eine Beispielaufgabe. Das Verhalten der Testteilnehmer in dieser Aufgabe wurde nicht in die Bewertung einbezogen. Für jede Aufgabe wurde ein Verkehrsszenario entwickelt und mithilfe des VICOM-Editors, dem Autorenprogramm, das auch in Vorstudie 1 zur computerbasierten Erstellung von bewegtem und unbewegtem Bildmaterial eingesetzt wurde, in eine bildliche Darstellungsform gebracht. Von jeder einzelnen Aufgabe wurde eine dynamische Version, in der eine Computeranimation präsentiert wurde, sowie eine statische Version mit einem Bild erstellt. Den Teilnehmern wurden entweder Computeranimationen der 32 Szenarien mit einer Länge von 19 bis 45 Sekunden (dynamische Version) oder Standbilder des entsprechenden Szenarios (statische Version) aus der Ego-Perspektive gezeigt. Die Probanden wurden instruiert, bei jedem Szenario so schnell wie möglich einmal die Leertaste zu drücken, sobald sie einen Hinweis darauf erkennen konnten, dass das Ego-Fahrzeug seine Geschwindigkeit reduzieren sollte. Dieses Aufgabenformat ähnelt damit einigen *Hazard-Perception*-Tests, die in Kapitel 2.3.4 beschrieben wurden. Die Videos bzw. Bilder wurden nacheinander in randomisierter Reihenfolge gezeigt. In der Instruktion wurden die Teilnehmer darüber in Kenntnis gesetzt, dass innerhalb jedes Videos maximal einmal die Notwendigkeit zur Geschwindigkeitsreduktion gegeben war. Diese Restriktion war notwendig, um das dynamische Format an das statische Format anzupassen und auf diese Art ihre Vergleichbarkeit zu gewährleisten.

Bei näherer inhaltlicher Betrachtung der einzelnen Szenarien können drei Aufgabentypen unterschieden werden. Aufgaben, die keinerlei Bremshinweise beinhalten, werden im Folgenden *Distraktoren* genannt. Aufgaben mit explizitem Bremshinweis werden als *Hazard-Perception*-Aufgaben, Aufgaben mit implizitem Bremshinweis werden als *Vorschriftaufgaben* bezeichnet. In Tabelle 4.11 erfolgt eine Auflistung der eingesetzten Aufgaben mit einer knappen Beschreibung ihrer Inhalte und ihrer Zuordnung zum jeweiligen Aufgabentyp. Als *Hazard-Perception*-Aufgaben waren insgesamt 19 Aufgaben zu bezeichnen. Innerhalb dieser Aufgaben tauchten demnach Bremsgründe auf, die als tatsächliche Gefahrenhinweise zu werten waren (z.B. ein Kind läuft plötzlich auf die Straße). Abbildung 4.9 zeigt ein solches *Hazard-Perception*-Szenario in seiner statischen Version. Innerhalb dieses Szenarios verfolgt der Fahrer des Ego-Fahrzeuges offensichtlich die Absicht,

den vorausfahrenden LKW zu überholen und blinkt links, um anzuzeigen, dass er mit relativ hoher Geschwindigkeit von rechts auf die mittlere Autobahnspur wechseln möchte. Gleichzeitig wechselt ein Fahrzeug von der linken Spur auf die mittlere Spur, sodass das Ego-Fahrzeug, um eine Kollision zu vermeiden, den Überholvorgang abbrechen und seine Geschwindigkeit reduzieren müsste. Um bei dieser Aufgabe rechtzeitig zu reagieren, bedarf es keines Wissens bezüglich Verkehrsregeln.

Zudem gab es sechs Vorschriftaufgaben, die einen eher impliziten Bremsgrund beinhalteten, der aus den in der Situation geltenden Verkehrsregeln abzuleiten war. In diesen Szenarien tauchten Hinweise darauf auf, dass bestimmte Verkehrsregeln (z.B. „Vorfahrt achten“) eine Geschwindigkeitsreduktion erforderten.



Abbildung 4.9
Statische Version der Reaktionszeitaufgabe Nr. 28 (Hazard Perception) (aus Malone et al., 2012; S. 45)

Abbildung 4.10 zeigt ein Bild das zu einer Vorschriftaufgabe gezeigt wurde und auf das in der statischen Version des Testmaterials reagiert werden sollte. Das Ego-Fahrzeug nähert sich dabei einer Einmündung, an der man den von rechts kommenden Fahrzeugen die Vorfahrt gewähren muss. Da man nicht weit in die Straße hineinschauen kann, um zu überprüfen ob ein Fahrzeug von rechts naht, ist eine Geschwindigkeitsreduktion notwendig.



Abbildung 4.10
Statische Version der Reaktionszeitaufgabe Nr. 15 (Vorschriftaufgabe)

Tabelle 4.11
Beschreibung der Reaktionszeitaufgaben

Nr.	Aufgabentyp	Beschreibung
1	Übungsaufgabe	Städtisch, Abbieger bremst unerwartet (ohne Anzeige des Abbiegevorgangs)
2	Hazard Perception	Verschneite Landstraße, entgegenkommender LKW auf eigener Spur
3	Hazard Perception	Schnee, kurvige Landstraße, vorausfahrendes Fahrzeug bremst
4	Hazard Perception	Innerorts, Fußgänger läuft über die Straße
5	Distraktor	Landstraße
6	Distraktor	Autobahntunnel
7	Distraktor	Autobahn
8	Distraktor	Landstraße
9	Vorschriftaufgabe	Landstraße, halb verdecktes Ortsschild bei 80 km/h
10	Hazard Perception	Innerorts, PKW schert plötzlich knapp von rechts ein
11	Hazard Perception	Städtisch, parkende PKWs und entgegenkommendes Fahrzeug
12	Hazard Perception	Innerorts, Kind läuft über die Straße
13	Hazard Perception	Innerorts, parkende PKWs am Fahrbahnrand
14	Hazard Perception	Schnee, Landstraße, enge Kurve mit 80 km/h aktueller Fahrgeschwindigkeit
15	Vorschriftaufgabe	Innerorts, Einmündung, rechts vor links
16	Distraktor	Innerorts
17	Vorschriftaufgabe	Landstraße, Bus an Haltestelle mit Warnblinklicht
18	Vorschriftaufgabe	innerorts, nachts, Vorfahrt gewähren an abknickender Vorfahrtsstraße
19	Vorschriftaufgabe	Autobahn, Regen, Schild „80 bei Nässe“ (bei aktuell 100 km/h)
20	Vorschriftaufgabe	Autobahn, rechts Überholen von langsamerem PKW
21	Vorschriftaufgabe	Innerorts, Schild „Kreuzung mit rechts vor links“
22	Distraktor	Innerorts
23	Hazard Perception	Landstraße, Ampel springt auf gelb
24	Hazard Perception	Innerorts, Hund läuft auf die Straße
25	Hazard Perception	Innerorts, Unimog schert an Baustelle auf die Fahrbahn aus
26	Hazard Perception	Innerorts, eigene Spur wird plötzlich blockiert
27	Hazard Perception	Innerorts, vorausfahrender PKW bremst wegen Fahrradfahrer
28	Hazard Perception	Autobahn, PKW auf der linken Spur wechselt gleichzeitig mit eigenem PKW auf die mittlere Spur
29	Hazard Perception	Innerorts, vorausfahrender Abbieger bremst wegen Fahrradfahrer
30	Hazard Perception	Innerorts, vorausfahrender LKW bremst plötzlich zum Abbiegen an Tankstelle
31	Hazard Perception	Autobahn, vorausfahrender LKW wechselt plötzlich auf die eigene Spur
32	Hazard Perception	Innerorts, PKW wechselt knapp auf die eigene Spur

Sechs weitere Aufgaben konnten als Distraktoren bezeichnet werden, da in diesen Animationen beziehungsweise auf diesen Standbildern kein Hinweis darauf zu erkennen war, dass man als Fahrer die Geschwindigkeit reduzieren sollte.

Die Erstellung der einzelnen Szenarien und die Auswahl relevanter Situationen erfolgten nach unterschiedlichen Gesichtspunkten. Elf der *Hazard-Perception*-Aufgaben wurden aus Aufgaben eines *Hazard-Perception*-Tests abgeleitet, der innerhalb des Forschungsprojektes *Wirkungsanalyse und Bewertung der neuen Regelungen im Rahmen der Fahrerlaubnis auf Probe* (Debus et al., 2008) der Bundesanstalt für Straßenwesen erstellt wurde, abgeleitet. Im Zuge dieses Projektes wurden den Teilnehmern, die über Fahrerfahrung in unterschiedlichem Ausmaß verfügten, reale Straßenverkehrsszenen auf Fotos dargeboten. Die Probanden sollten dabei reagieren, sobald sie auf einem Foto, das im Realverkehr aufgenommen wurde, einen Gefahrenhinweis entdeckten. Die Daten aus diesem Forschungsprojekt wurden auf der Aufgabenebene analysiert und somit konnten Fotos identifiziert werden, auf die Novizen signifikant langsamer reagiert hatten als Experten (Aufgaben Nr. 2, 3, 4, 9, 10, 11, 12 und 23). Es wurden einige dieser Szenarien für die vorliegende Arbeit ausgewählt, die sich bereits bewährt hatten, da es für diese Aufgaben als empirisch gesichert gelten kann, dass es sich bei den dargestellten Szenarien um sicherheitsrelevante Verkehrssituationen handelte. Um zu überprüfen, ob sich die Aufgaben bei dynamischer Darbietung der Szenarien ebenfalls als kriteriumsvalide erweisen oder sogar noch deutlicher zwischen Könnern und Nicht-Könnern trennen, wurden diese Szenarien in die Studie einbezogen. Des Weiteren wurden Szenarien einbezogen, in denen Experten keine besseren Leistungen als Novizen erbracht hatten (Aufgaben Nr. 13, 14 und 15). Durch die vorliegende Studie sollte ermittelt werden, ob diese durch den Einbezug dynamischer Darstellung in ihrer Trennschärfe verbessert werden können. Die aus der Studie von Debus et al. (2008) entliehenen Realfotos wurden dazu zum einen für die statische Version in computererstellte Bilder umgewandelt und zum anderen wurden für die dynamische Version kurze Animationen erstellt, die ggf. mit einer kritischen Situation endeten.

Zehn weitere Szenarien stammen aus einem an der TU Chemnitz entwickelten Lernprogramm für Fahrschüler (Aufgaben Nr. 24, 25, 26, 27, 28, 29, 30, 31 und 32), das im Rahmen eines parallel zur vorliegenden Studie laufenden Forschungsprojektes im Auftrag der Bundesanstalt für Straßenwesen getestet wurde (Petzoldt, Weiß, Franke, Krems & Bannert, 2011). Die Verkehrssituationen, die für dieses Lernprogramm ausgewählt wurden, stellen Verkehrsszenarien dar, die theoretisch und empirisch begründete Sicherheitsrisiken für

Fahranfänger bergen. Die dynamischen Darstellungen wurden an der TU Chemnitz mit dem VICOM-Editor erstellt und mussten nur geringfügig an das Testverfahren angepasst werden. Eine entsprechende statische Version jedes Szenarios musste zusätzlich erstellt werden.

Vier Szenarien der Vorschriftaufgaben wurden aus Aufgaben der theoretischen Fahrerlaubnisprüfung abgeleitet (Aufgaben Nr. 17, 18, 19 und 20), die in ihrer Originalversion durch ein Foto oder eine Zeichnung angereichert waren. Ausgehend von der Situationsbeschreibung und dem Foto, wurde die Aufgabe in einer statischen und einer dynamischen Version, angepasst an das Aufgabenformat, erstellt.

Persönliche Angaben

Es wurden analog zur Vorgehensweise in Vorstudie 1 Angaben zu demographischen Daten und zum Verkehrsverhalten und Unfallgeschichte erfragt (vgl. Kapitel 4.1.2). In dieser Studie erfolgte die Erhebung allerdings in Papier-Bleistift-Form.

Sensation-Seeking-Skala

Die Tendenz zur Sensationssuche, die mit dem Risikoverhalten beim Fahren in Verbindung gebracht wird, wird in der Regel mit der *Sensation Seeking Scale* (SSS-V) nach Zuckerman (1979) erfasst (Jonah, 1996). Diese Skala besteht insgesamt aus vier Subskalen: *Thrill and Adventure Seeking* (TAS), *Experience Seeking* (ES), *Disinhibition* (DIS) und *Boredom Susceptibility* (BS). Den Versuchspersonen wurden lediglich 12 Items der TAS-Skala zur Beantwortung vorgegeben. Die Items waren dabei zwei Unterskalen zugeordnet. Der Skala „*Thrill and Adventure Seeking* mit vitalem Risiko“ (TAS-vr) gehörten acht Items an. Dabei handelte es sich um Aussagen bzgl. der Präferenzen für Erlebnisse, die körperliche Aktionen beinhalten, bei denen man einem tatsächlichen gesundheitlichen Risiko ausgesetzt ist (z.B. „Wenn ich die Möglichkeit hätte, würde ich bestimmt Fallschirmspringen oder Drachenfliegen“). Der Skala „*Thrill and Adventure Seeking* ohne vitales Risiko“ (TAS-ovr) wurden vier Items zugeordnet. Diese Skala soll die Tendenz zum Aufsuchen von Nervenkitzel ohne die Gefährdung der eigenen Person erfassen (z.B. „Ich lasse mich gerne von aufpeitschender, lauter Musik anheizen“). Die Items waren in Anlehnung an Gniech, Oetting und Brohl (1993) jeweils auf einer fünfstufigen Ratingskala mit den Endpunkten +2 („trifft zu“) und -2 („trifft nicht zu“) zu beantworten. Der Mittelpunkt der Skala war gekennzeichnet mit „unentschieden“, um den Versuchspersonen die Gelegenheit zu geben, auf Aussagen zu reagieren, bei denen sie weder eine bejahende noch eine verneinende Antwort geben wollten.

Zur Berechnung der individuellen Ausprägung des Merkmals wurden die Mittelwerte aus den zwölf Items, bzw. den Items der beiden Unterskalen, berechnet.

4.2.2.3 Durchführung

Die Anwerbung von Versuchspersonen erfolgte zum einen durch Aushänge an der Universität des Saarlandes. Zum anderen wurden Fahrschüler unterschiedlicher Fahrerlaubnisklassen in fünf saarländischen Fahrschulen im Theorieunterricht angeworben. Die Datenerhebung erfolgte entweder im Computerlabor der Fachrichtung Bildungswissenschaften auf dem Campus der Universität des Saarlandes oder mithilfe eines mobilen Computerlabors in den Räumen der Fahrschulen. Sowohl in den Fahrschulen als auch auf dem Universitätscampus nahmen Personen teil, die bereits über eine Fahrerlaubnis der Klasse B verfügten, und solche, die die Fahrausbildung noch nicht absolviert hatten. In der Regel erfolgte die Erhebung in Gruppen zwischen zwei und zehn Versuchsteilnehmern.

Die Teilnehmer bearbeiteten entweder die statische oder die dynamische Version des Versuchsmaterials. Die Zuweisung der Bedingungen des Faktors Präsentationsform (statisch vs. dynamisch) erfolgte zufällig. Den Versuchsteilnehmern wurde beim Eintreten abwechselnd ein Computer, auf dem die dynamische oder die statische Testversion eingestellt war, zugewiesen. Dabei wurde zusätzlich auf die Balance der Variablen Geschlecht und Expertise geachtet. Hinsichtlich des Faktors Expertise handelt es sich bei der vorliegenden Studie ebenso wie in Vorstudie 1 um ein Quasi-Experiment. Eine zufällige Einteilung der Versuchsteilnehmer in Novizen und Experten war nicht möglich, die beiden Gruppen bestanden bereits im Vorhinein. Die Versuchsleitung wurde von der Autorin selbst oder einer studentischen Hilfskraft der Fachrichtung Bildungswissenschaften der Universität des Saarlandes übernommen.

Nach einer kurzen Begrüßung wurde den Versuchsteilnehmern eine Aufwandentschädigung von fünf Euro gezahlt oder – bei Psychologiestudenten – 50 Versuchspersonenminuten bescheinigt. Die gesamte Erhebung dauerte etwa zehn Minuten länger, wenn die Teilnehmer die Reaktionszeitaufgaben in ihrer dynamischen Version bearbeiteten, als wenn ihnen die statische Version zur Bearbeitung vorlag. Insgesamt benötigten die Versuchsteilnehmer zur Bearbeitung des Testmaterials mit Animationen zwischen 40 und 50 Minuten und ohne Animationen zwischen 30 und 40 Minuten. Verständnisfragen, die während der Bearbeitung gestellt wurden, wurden von den Versuchsleitern beantwortet. Die Erhebung begann für jeden Versuchsteilnehmer mit der Bearbeitung des Papier-Bleistift-Fragebogens zu persönlichen Angaben und der *Sensation*

Seeking-Skala. Anschließend starteten die Teilnehmer selbstständig das Testprogramm am Computer. Zunächst wurden die Teilnehmer in einem Text am Bildschirm über den Hintergrund der Untersuchung aufgeklärt. Danach wurde ein Instruktionstext dargeboten, der die Teilnehmer darüber informierte, worin ihre Aufgabe bei der Bearbeitung der Reaktionszeitaufgaben bestand. Im Anschluss daran wurden die Versuchsteilnehmer entweder zur statischen Version oder zur dynamischen Version der Reaktionszeitaufgaben weitergeleitet. Die Reihenfolge der einzelnen Aufgaben war zufällig. Vor jeder einzelnen Aufgabe wurde eine Kurzinstruktion eingeblendet, um die Teilnehmer an ihre Aufgabe zu erinnern, bevor das Bild oder Video eingeblendet wurde. Jede Aufgabe wurde durch den Teilnehmer selbst gestartet. Über das Ende der Erhebung wurde durch eine Mitteilung am Bildschirm hingewiesen, nachdem die letzte Aufgabe bearbeitet wurde.

4.2.3 Ergebnisse

Für alle statistischen Tests, die im Folgenden berichtet werden, wurde ein Alphaniveau von .05 festgesetzt. Das Alphaniveau wurde für Einzelvergleiche durch das jeweils angegebene Verfahren angepasst.

4.2.3.1 Item- und Skalenstatistiken der abhängigen Variablen

Zur Aufgabenselektion und Einschätzung der Reliabilität der Aufgaben wurden Skalen- und Itemkennwerte berechnet. Analog zum Vorgehen in Vorstudie 1 wurden die Trennschärfen durch die korrigierten Item-Skala-Korrelationen definiert und zusätzlich wurden als Fremdtrennschärfen die Korrelationen zwischen den Leistungen in den einzelnen Aufgaben und dem Kriterium Expertise berechnet. In Anhang A werden die Kennwerte für die einzelnen Aufgaben berichtet, während in Tabelle 4.12 ein Überblick gewährt wird.

Die Reliabilität der Reaktionszeitaufgaben wird getrennt für die Lösung der Aufgaben (korrekt oder falsch gelöst) und für die Reaktionszeiten (z-Werte) berichtet. Die Reliabilität der Gesamtskala (statisch und dynamisch) für die Lösung der Aufgaben ist mit einem Alphaswert (nach Kuder Richardson Formula 20) von .77 als zufriedenstellend zu bezeichnen. Etwas geringer waren die internen Konsistenzen getrennt berechnet für die statische ($\alpha = .72$) und die dynamische Version der Aufgaben ($\alpha = .58$).

Die Distraktoraufgaben, auf die nicht reagiert werden sollte, gingen nicht in die Analyse der Reaktionszeiten ein. Wenn Personen auf einen Bremshinweis nicht reagiert hatten, wurden die fehlenden Werte für die Itemanalyse in der dynamischen Version durch die Zeiten ersetzt, die die Personen erreicht hätten, wenn Sie gerade noch rechtzeitig innerhalb

des kritischen Intervalls reagiert hätten. In der statischen Version wurden die fehlenden Werte durch die maximale Reaktionszeit in der jeweiligen Aufgabe ersetzt, die die anderen Teilnehmer gezeigt hatten. Zur Berechnung der internen Konsistenz sowie der Itemkennwerte der Reaktionszeiten wurden z-Werte gebildet. Dabei wurden die einzelnen Reaktionszeiten für jede Aufgabe am gemeinsamen Mittelwert der statischen und der dynamischen Version dieser Aufgabe standardisiert. Die Reliabilität der Reaktionszeiten war für die statische und die dynamische Versionen insgesamt mit einer internen Konsistenz von Cronbachs α von .82 als zufriedenstellend zu bezeichnen. Für die Versionen getrennt zeigten sich geringere Trennschärfen ($\alpha_{\text{statisch}} = .69$; $\alpha_{\text{dynamisch}} = .49$). Die internen Konsistenzen der beiden Skalen Lösung der Aufgaben und Reaktionszeiten weisen insgesamt auf eine zufriedenstellende Reliabilität hin. Die einzelnen Aufgaben können zu einer Skala zusammengefasst werden. Es zeigen sich allerdings deutlich geringere interne Konsistenzen für beide Skalen, wenn man die dynamische Version der Aufgaben separat betrachtet.

Für eine Testendform, die in die Längsschnittuntersuchung eingehen sollte, wurden Aufgaben, die geringe Trennschärfen und gleichzeitig geringe Fremdtrennschärfen aufwiesen ausgeschlossen. Die Schwierigkeit der Aufgaben soll allerdings kein Ausschlusskriterium darstellen. Aufgabe Nr. 20 sollte nicht in eine Testendform eingehen, da diese Aufgabe bezüglich der richtigen Lösungen und Reaktionszeiten niedrige Trennschärfen und Fremdtrennschärfen aufweist.

Mit einem MSA von .69 erfüllen die Daten die Voraussetzungen zur Faktorenanalyse. Es wurde eine Hauptkomponentenanalyse durchgeführt. Der Eigenwerteverlauf bestätigte die einfaktorielle Struktur der Daten zwar, der extrahierte Faktor erklärte allerdings lediglich 17 % der Varianz.

Zur Ermittlung der *Sensation-Seeking*-Tendenz der Teilnehmer wurden ihnen 12 Items der TAS-Skala zur Bearbeitung vorgelegt. Aus ihren Antworten zu jeder Aussage auf einer fünfstufigen Skala wurde der Mittelwert aller Aufgaben berechnet. In Tabelle 4.13 werden die deskriptiven Kennwerte der Novizen und Experten bezüglich der Gesamtskala TAS und den beiden Subskalen TAS-ovr und TAS-vr dargestellt. Zur Ermittlung der Unterschiede zwischen den Experten und den Novizen, hinsichtlich ihrer *Sensation-Seeking*-Tendenzen in der Gesamtskala und den beiden Subskalen, wurde eine Varianzanalyse mit dem unabhängigen Faktor Expertise (Experten vs. Novizen) und den abhängigen Variablen *Thrill and Adventure Seeking* (TAS), TAS ohne vitales Risiko (TAS-ovr) und TAS mit vitalem Risiko (TAS-vr) durchgeführt.

Tabelle 4.12
Itemkennwerte der Reaktionszeitaufgaben (Übersicht)

	Gesamt N = 136	Statisch n = 63	Dynamisch n = 73
Mittlere Schwierigkeit (<i>SD</i>)	.64 (.44)	.75 (.37)	.55 (.45)
Schwierigkeit (Min - Max)	.22 bis .90	.26 bis .98	.16 bis .90
Mittlere Trennschärfe (Lösung)	.28	.26	.17
Trennschärfe (Lösung) (Min - Max)	-.03 bis .57	-.14 bis .51	-.05 bis .41
Mittlere Fremdtrennschärfe (Lösung)	.28	.18	.11
Fremdtrennschärfe (Lösung) (Min - Max)	-.11 bis .36	.00 bis .51**	-.18 bis .38**
Mittlere Fremdtrennschärfe (z-Werte)	.36	.32	.13
Fremdtrennschärfe (z-Werte) (Min - Max)	-.05 bis .66	-.07 bis .56	-.08 bis .41
Mittlere Fremdtrennschärfe (z-Werte)	-.08	-.06	-.18
Fremdtrennschärfe (z-Werte) (Min - Max)	-.24* bis .24	-.28* bis .31	-.66** bis .67*

*Anmerkungen. SD = Standardabweichung. * Die Korrelation ist auf dem Niveau von .05 (2-seitig) signifikant.
 ** Die Korrelation ist auf dem Niveau von .01 (2-seitig) signifikant.*

Es zeigten sich keine statistisch bedeutsamen Unterschiede zwischen Experten und Novizen in der Gesamtskala ($F(1; 139) = 1.55; p = .22$). Die Experten und die Novizen unterschieden sich auch nicht bzgl. ihrer Angaben in der Subskala TAS-ovr ($F(1, 139) = .79; p = .38$) und in der Subskala TAS-vr ($F(1, 139) = 1.26; p = .26$).

Statistisch bedeutsame Korrelationen zwischen der Leistung in den Reaktionszeitaufgaben (Anzahl der korrekten Lösungen, Reaktionszeiten) und den Werten der Teilnehmer in der Gesamtskala und den beiden Subskalen zeigten sich nicht. Lediglich für die Novizen zeigte sich eine signifikante negative Korrelation zwischen den Reaktionszeiten (z-

Werte) und den Werten in der Skala TAS-vr ($r = -.29$; $p = .03$). Je deutlicher den Aussagen dieser Skala zugestimmt wurde, desto schneller wurde auf Bremshinweise reagiert.

Tabelle 4.13

Ergebnisse in der Kontrollvariablen Sensationssuche: Mittelwerte und Standardabweichungen in den TAS-Skalen

	TAS gesamt		TAS-ovr		TAS-vr	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Novizen n = 61	.37	.45	.39	.59	.36	.52
Experten n = 79	.47	.48	.47	.57	.46	.58
Gesamt N = 140	.42	.47	.44	.58	.41	.55

Anmerkungen. *M* = Mittelwert. *SD* = Standardabweichung

4.2.3.2 Einfluss von Expertise und Präsentationsform auf die Leistung in den Reaktionsaufgaben

Um den Einflusses von Fahrexpertise und Präsentationsform auf die Leistung in den Reaktionsaufgaben zu ermitteln, wurden die Daten mithilfe einer Varianzanalyse mit Messwiederholung auf einem Faktor analysiert. Die Faktoren Expertise (Experten vs. Novizen) und Präsentationsform (statisch vs. dynamisch) gingen als unabhängigen Variablen in die Analyse ein. Als Messwiederholungsfaktor ging der dreifach gestufte Faktor Aufgabentyp (*Hazard Perception* vs. *Vorschrift* vs. *Distraktor*) in die Varianzanalyse ein. Abhängige Variablen waren die mittleren Lösungswahrscheinlichkeiten der Items und die z-standardisierten mittleren Reaktionszeiten. Für die beiden abhängigen Variablen wurden zwei getrennte Varianzanalysen durchgeführt, da im Falle der Reaktionszeiten eine Faktorstufe wegfällt: Für Distraktoren wurden keine Reaktionszeiten erhoben.

Aufgrund der fehlenden Gruppenunterschiede und dem geringen Zusammenhang zwischen den *Sensation-Seeking*-Skalen und den Leistungswerten der Reaktionszeitaufgaben, ging die Variable *Sensation Seeking* nicht als Kovariate in die Varianzanalysen ein.

Abhängige Variable: Lösungswahrscheinlichkeit

In Tabelle 4.14 werden die deskriptiven Kennwerte der Leistungen der Teilnehmer für die abhängige Variable mittlere Lösungswahrscheinlichkeit dargestellt.

Es zeigte sich ein Haupteffekt der Expertise ($F(1, 132) = 33.21, p < .001; \eta_p^2 = .20$). Der Vergleich der Mittelwerte verdeutlicht, dass Experten ($M = .70; SD = .12$) im Mittel bessere Leistungen erbrachten als die Novizen ($M = .60; SD = .15$); Sie lösten insgesamt mehr Aufgaben richtig. Zudem zeigte sich ein Haupteffekt der Präsentationsform ($F(1, 132) = 55.67, p < .001, \eta_p^2 = .30$). Die Analyse der Mittelwerte zeigte, dass die Aufgaben in der statischen Version des Testmaterials ($M = .73; SD = .12$) insgesamt häufiger gelöst wurden als in der dynamischen Testversion ($M = .59; SD = .13$). Die Interaktion der beiden Faktoren Expertise und Präsentationsform hielt einer Prüfung auf statistische Signifikanz nicht stand ($F(1, 132) = .05; p = .82$).

Die Varianzanalyse zeigte einen signifikanten Haupteffekt des Messwiederholungsfaktors Aufgabentyp ($F(2, 264) = 108.16; p < .001; \eta_p^2 = .45$). Wie in Abbildung 4.11 veranschaulicht wird, sinkt die Lösungswahrscheinlichkeit von Distraktoren über *Hazard-Perception*-Aufgaben zu Vorschriftaufgaben. Dies verdeutlicht auch den Unterschied der beiden Fehlerarten: Die Wahrscheinlichkeit dafür, dass reagiert wurde, obwohl kein relevanter Reiz vorhanden war (false positives) war geringer als die Wahrscheinlichkeit einen relevanten Reiz nicht zu entdecken (misses). Dabei waren die Differenzen zwischen allen drei Aufgabentypen statistisch bedeutsam (alle $p < .001$). Für die Paarvergleiche wurde eine Anpassung des α -Niveaus nach Bonferroni vorgenommen.

Ein Interaktionseffekt zwischen Aufgabentyp und Expertise konnte nicht gezeigt werden ($F(2, 264) = 1.39; p = .25$). Allerdings konnte eine signifikante Interaktion zwischen den Faktoren Aufgabentyp und Präsentationsform festgestellt werden ($F(2, 264) = 23.02; p < .001; \eta_p^2 = .15$), die in Abbildung 4.12 veranschaulicht wird. Um die Interaktion genauer zu spezifizieren, wurden Scheffé-Tests für die Einzelvergleiche durchgeführt. Dabei zeigten sich statistisch bedeutsame Unterschiede zwischen den mittleren Leistungen in den beiden Testversionen lediglich bei den Distraktor-Aufgaben ($D = .13; \text{Diff}_{\text{krit}} = .11; p < .05$) und den *Hazard-Perception*-Aufgaben ($D = .30; \text{Diff}_{\text{krit}} = .16; p < .001$).

Tabelle 4.14

Ergebnisse in den Reaktionszeitaufgaben (mittlere Lösungswahrscheinlichkeiten): Mittelwerte und Standardabweichungen der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch)

Aufgabentyp	Expertise	Präsentationsform	M	SD	n
Distraktor	Novizen	statisch	.82	.18	29
		dynamisch	.69	.25	29
		gesamt	.75	.22	58
	Experten	statisch	.95	.09	34
		dynamisch	.81	.22	44
		gesamt	.87	.19	78
	gesamt	statisch	.89	.15	63
		dynamisch	.76	.24	73
		gesamt	.82	.21	136
Hazard Perception	Novizen	statisch	.73	.18	29
		dynamisch	.48	.14	29
		gesamt	.61	.20	58
	Experten	statisch	.85	.10	34
		dynamisch	.51	.14	44
		gesamt	.66	.21	78
	gesamt	statisch	.80	.15	63
		dynamisch	.50	.14	73
		gesamt	.64	.21	136
Vorschrift	Novizen	statisch	.46	.25	29
		dynamisch	.41	.21	29
		gesamt	.44	.23	58
	Experten	statisch	.57	.20	34
		dynamisch	.59	.16	44
		gesamt	.58	.18	78
	Gesamt	statisch	.52	.23	63
		dynamisch	.52	.20	73
		gesamt	.52	.21	136

Anmerkungen. M = Mittelwert. SD = Standardabweichung

In der statischen Version wurden bessere Leistungen erbracht als in der dynamischen Version des Testmaterials. In den Vorschrift-Aufgaben unterschieden sich die Leistungen, die in den beiden Testversionen erreicht wurden, nicht in statistisch bedeutsamem Ausmaß ($D = 0$; $\text{Diff}_{\text{krit}} = .11$; $p > .05$).

Die dreifache Wechselwirkung zwischen den Faktoren Aufgabentyp, Präsentationsform und Expertise hielt einer Überprüfung auf statistische Signifikanz nicht stand ($F(2, 264) = 1.66$; $p = .19$).

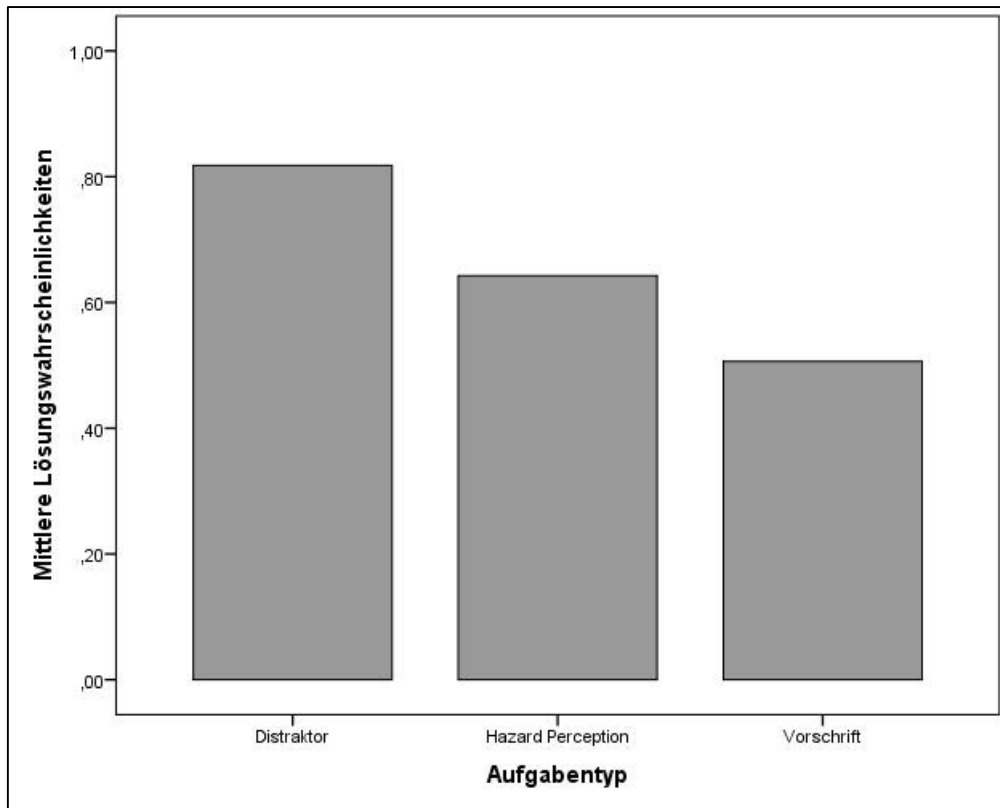


Abbildung 4.11
 Ergebnisse zu Reaktionszeitaufgaben (Lösungswahrscheinlichkeit): Haupteffekt Aufgabentyp

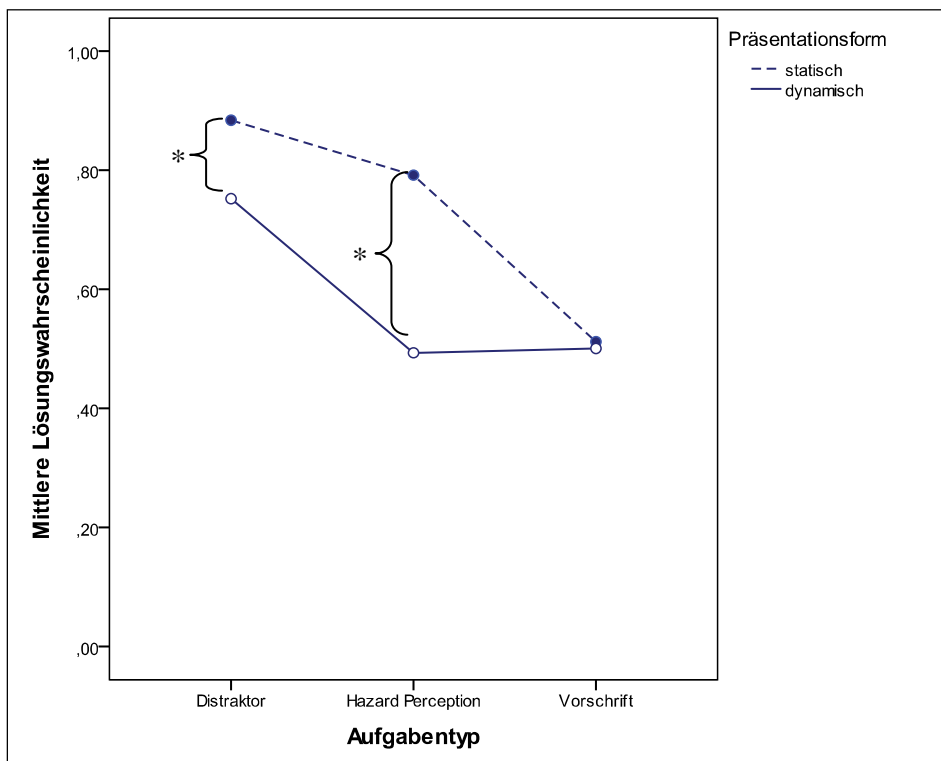


Abbildung 4.12
 Ergebnisse zu Reaktionszeitaufgaben (Lösungswahrscheinlichkeit): Interaktion Aufgabentyp*Präsentationsform

Als abhängige Variable wurde in den Analysen für die Leistung der Probanden die mittlere Lösungswahrscheinlichkeit der Aufgaben eingesetzt. Antworttendenzen wurden dabei nicht in die Leistungsmessung einbezogen. Ein anderer Ansatz zur Ermittlung der Leistung der Teilnehmer ergibt sich nach der Signalentdeckungstheorie (Swets, 1961). Analog zum Vorgehen bei Studien zur Signalentdeckungstheorie wurde als Leistungskriterium d' berechnet. Dazu wurden die Summen der richtigen Reaktionen (Reaktion bei Vorschrift- und *Hazard-Perception*-Aufgaben = *hits*) und der falschen Reaktionen (Reaktion bei Distraktoren = *false alarms*) an den jeweiligen Mittelwerten z-standardisiert. Durch die Subtraktion des z-Wertes für die falschen Reaktionen vom z-Wert für die richtigen Reaktionen wurde d' berechnet. Eine Varianzanalyse mit den beiden unabhängigen Faktoren Expertise und Präsentationsform und der abhängigen Variable d' zeigte einen signifikanten Haupteffekt des Faktors Expertise ($F(1, 132) = 32.46; p < .001; \eta_p^2 = .20$). Die Experten ($M = .49; SD = 1.39$) zeigten einen höheren Anteil richtiger Lösungen als die Novizen ($M = -.66; SD = 1.62$). Zudem zeigte sich ein signifikanter Haupteffekt der Präsentationsform ($F(1, 132) = 51.19; p < .001; \eta_p^2 = .27$). In der statischen Version des Testmaterials ($M = .81; SD = 1.22$) wurden relativ mehr richtige Reaktionen gezeigt als in der dynamischen Version des Testmaterials ($M = -.70; SD = 1.55$). Ein Interaktionseffekt zwischen den Faktoren Expertise und Präsentationsform konnte nicht festgestellt werden ($F(1, 132) = .04; p = .85$).

Abhängige Variable: Reaktionszeiten

Die Ermittlung der Reaktionszeit auf einen Reiz, der zu einer Bremsreaktion führen sollte, erfolgte in der dynamischen Version des Testmaterials durch die Aufzeichnung der Zeit zwischen dem ersten Auftauchen des kritischen Reizes im Video und der ersten Reaktion des Versuchsteilnehmers durch Betätigung der Leertaste. Als Reaktionszeit in der statischen Version wurde die Zeit erfasst, die vom Erscheinen des Bildes bis zur Reaktion durch die Versuchsperson verstrich. Die Reaktionszeit einer Person bei einer bestimmten Aufgabe wurde zudem nur in die Analyse einbezogen, wenn auf das Bild überhaupt reagiert wurde (statische Version), bzw. wenn innerhalb der Animation im zuvor festgelegten kritischen Zeitfenster (erstes Auftauchen des kritischen Reizes bis Ende der Animation) reagiert wurde (dynamische Version). Die Reaktionszeiten auf die Distraktoren wurden nicht einberechnet, da hier eine Reaktion als Falschantwort gewertet wurde. Um einen Vergleich der Reaktionszeiten zwischen den beiden Präsentationsformaten (statisch vs. dynamisch) zu ermöglichen, erfolgte eine z-Standardisierung der Werte am gemeinsamen Mittelwert der statischen und der dynamischen Version jeder Aufgabe. In Tabelle 4.15 werden die

deskriptiven Kennwerte getrennt nach Expertise, Präsentationsform und Aufgabentyp dargestellt.

Eine Varianzanalyse mit den Faktoren Expertise (Experten vs. Novizen) und Präsentationsform (statisch vs. dynamisch) und dem zweifach gestuften Messwiederholungsfaktor Aufgabentyp (*Hazard Perception* vs. Vorschrift) zeigte einen tendenziellen Haupteffekt der Expertise ($F(1, 130) = 3.56$; $p = .06$; $\eta_p^2 = .03$). Die Novizen ($z = .06$; $SD = .66$) reagierten im Durchschnitt langsamer als die Experten ($z = -.11$; $SD = .43$) auf Reize, die eine Notwendigkeit zum Bremsen anzeigten.

Tabelle 4.15

Reaktionszeitaufgaben (z-Werte): Mittelwerte und Standardabweichungen der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch)

Aufgabentyp	Expertise	Präsentationsform	<i>M</i>	<i>SD</i>	<i>n</i>
Hazard Perception	Novizen	statisch	.21	.84	29
		dynamisch	-.07	.33	27
		gesamt	.07	.66	56
	Experten	statisch	-.02	.55	34
		dynamisch	-.20	.26	44
		gesamt	-.12	.42	78
	gesamt	statisch	.09	.70	63
		dynamisch	-.15	.30	71
		gesamt	-.04	.54	134
Vorschrift	Novizen	statisch	-.19	.88	29
		dynamisch	.35	.56	27
		gesamt	.07	.79	56
	Experten	statisch	-.15	.78	34
		dynamisch	-.06	.30	44
		gesamt	-.10	.56	78
	gesamt	statisch	-.17	.82	63
		dynamisch	.10	.46	71
		gesamt	-.03	.67	134

Anmerkungen. *M* = Mittelwert. *SD* = Standardabweichung

Deskriptiv zeigte sich, dass in der statischen Version des Testmaterials ($z = -.04$; $SD = .71$) etwas schneller reagiert wurde als in der dynamischen Version ($z = -.03$; $SD = .33$). Es konnte jedoch weder ein Haupteffekt Präsentationsform ($F(1, 130) = .18$; $p = .67$) noch eine Interaktion zwischen den Faktoren Expertise und Präsentationsform ($F(1, 130) = 0.88$; $p = .35$) festgestellt werden.

Auch ein Haupteffekt des Messwiederholungsfaktors Aufgabentyp konnte nicht nachgewiesen werden ($F(1, 130) = .06$; $p = .80$). Eine Wechselwirkung zwischen dem Aufgabentyp und dem Faktor Expertise lag ebenfalls nicht vor ($F(1, 130) < .001$; $p = .96$). Es

zeigte sich eine statistisch bedeutsame Interaktion zwischen den Faktoren Aufgabentyp und Präsentationsform ($F(1, 130) = 46.94$; $p < .001$; $\eta_p^2 = .27$). Die Interaktion zwischen den Faktoren Aufgabentyp und Präsentationsformat wird in Abbildung 4.13 veranschaulicht.

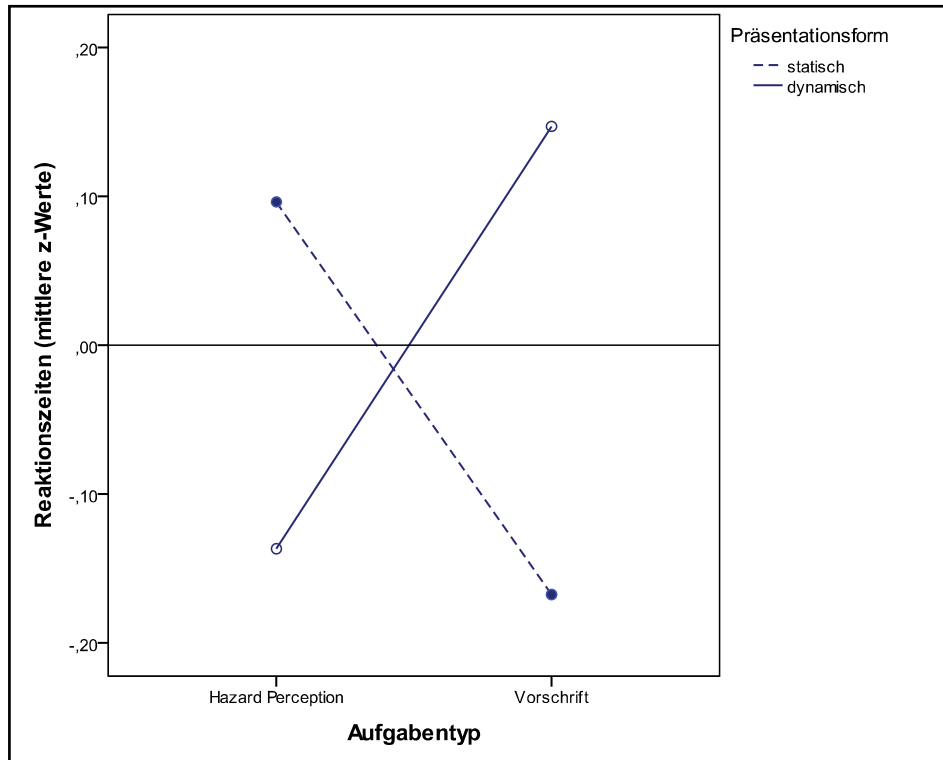


Abbildung 4.13

*Ergebnisse zu Reaktionszeitaufgaben (z-Werte): Interaktion Aufgabentyp*Präsentationsform*

Die mittleren z-Werte der Reaktionszeiten für die einzelnen Bedingungen zeigen, dass bei Aufgaben des Aufgabentyps *Hazard Perception* in der statischen Version langsamer reagiert wurde als in der dynamischen Version, während sich beim Aufgabentyp *Vorschrift* ein umgekehrtes Bild zeigte. Der Scheffé-Test bestätigte eine statistisch bedeutsame Differenz zwischen den Zeiten in der statischen und der dynamischen Version für die *Hazard-Perception*-Aufgaben ($D = .24$; $\text{Diff}_{\text{krit}} = .18$; $p < .001$) sowie für die *Vorschrift*-aufgaben ($D = .26$; $\text{Diff}_{\text{krit}} = .18$; $p < .001$). Ein weiteres Ergebnis der Varianzanalyse war eine dreifache Interaktion der Faktoren Aufgabentyp, Expertise und Präsentationsform ($F(1, 130) = 11.47$; $p = .001$; $\eta_p^2 = .08$). Die dreifache Wechselwirkung wird in Abbildung 4.14 veranschaulicht.

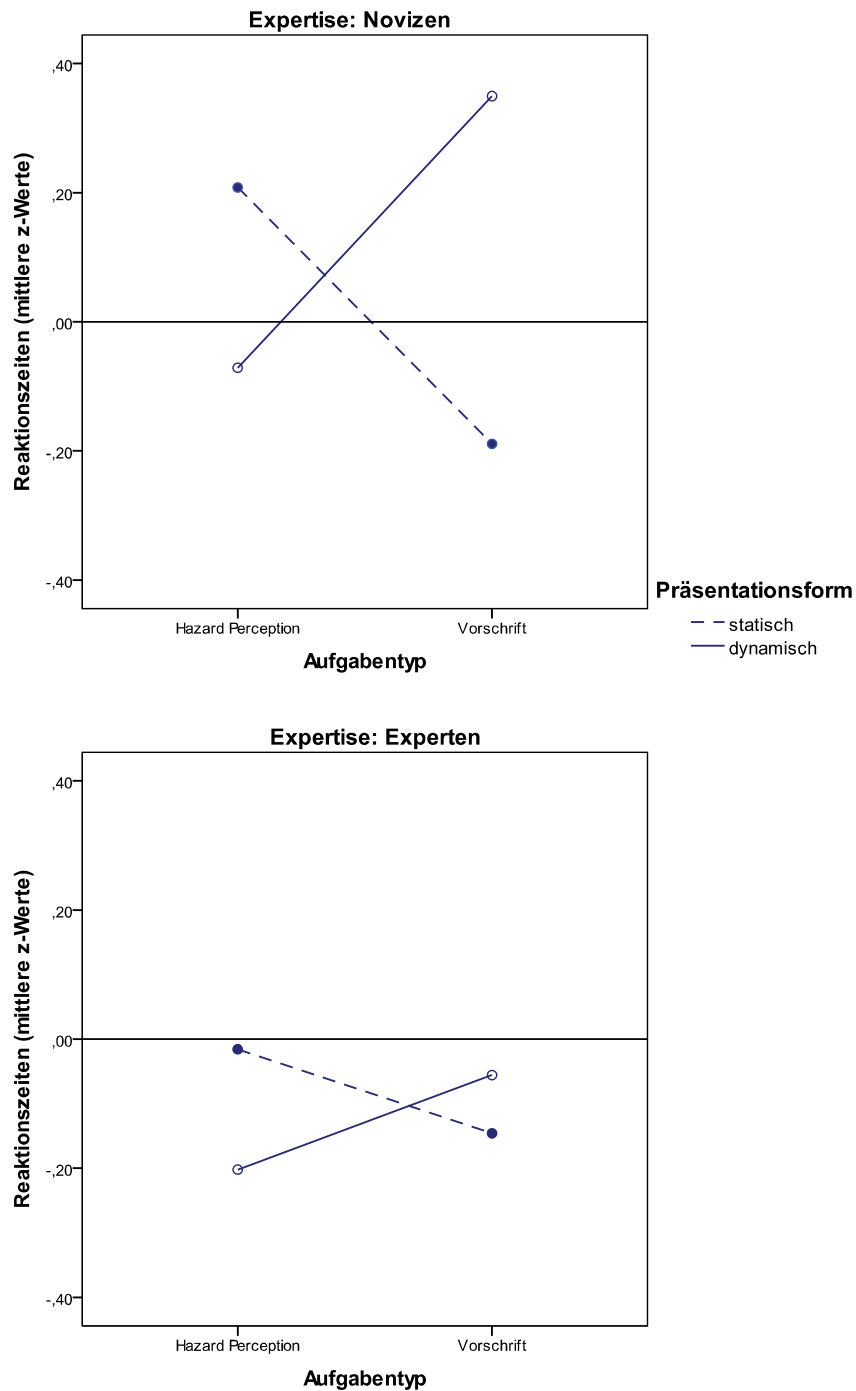


Abbildung 4.14
 Ergebnisse zu Reaktionszeitaufgaben (z-Werte): Interaktion
 Aufgabentyp*Präsentationsform*Expertise

Es zeigte sich, dass die Interaktion zwischen den Faktoren Aufgabentyp und Präsentationsform durch den dritten Faktor Expertise moderiert wird.

Einzelvergleiche mit dem Scheffé-Test zeigten für die Novizen für die *Hazard-Perception*-Aufgaben keine signifikanten Unterschiede zwischen der Leistung in der statischen und der dynamische Version des Testmaterials ($D = .27$; $\text{Diff}_{\text{krit}} = .58$; $p > .05$). Für

die Vorschriftaufgaben waren deutlichere Unterschiede zwischen den beiden Testversionen festzustellen. Im konservativen Scheffé-Test hält dieser Unterschied einer Überprüfung der Signifikanz jedoch nicht stand ($D = .54$; $\text{Diff}_{\text{krit}} = .58$; $p > .05$).

Für die Experten zeigte sich ein umgekehrtes Bild. Für die *Hazard-Perception*-Aufgaben ($D = .18$; $\text{Diff}_{\text{krit}} = .49$; $p > .05$) ist der Unterschied zwischen den Testversionen zwar deutlicher als für die Vorschriftaufgaben ($D = .09$; $\text{Diff}_{\text{krit}} = .49$; $p > .05$), jedoch sind die Unterschiede für beide Aufgabentypen nicht statistisch bedeutsam.

4.2.4 Diskussion

Die nachfolgende Diskussion beinhaltet zunächst die Zusammenfassung der Ergebnisse der zweiten Vorstudie und zudem eine Darstellung der Bedeutung der einzelnen Ergebnisse im Bezug auf die aufgestellten Hypothesen. In einem zweiten Teil der Diskussion werden Implikationen für das weitere empirische Vorgehen behandelt.

4.2.4.1 Bedeutung der Ergebnisse für die Hypothesen

Das Ziel der zweiten Vorstudie bestand darin, zu ermitteln, welchen Einfluss Fahrexpertise und Präsentationsformat auf die Leistung in einem Reaktionszeitaufgabenformat haben. Das Reaktionszeitaufgabenformat, das bereits in früheren Studien in *Hazard-Perception*-Tests zum Einsatz kam, sollte hinsichtlich seiner generellen Güte bewertet werden. Innerhalb des Formates sollte der Einfluss verschiedener inhaltlicher Aufgabenmerkmale betrachtet werden. Die Ergebnisse sollten als Grundlage zur Entscheidung dienen, inwieweit sich das Reaktionszeitformat generell zur Diagnostik eines Teilbereiches der Fahrexpertise eignet.

Entsprechend den Ergebnissen der Expertiseforschung generell sowie den Befunden zum Einsatz von Reaktionszeittests zur Messung der Gefahrenwahrnehmung, wurde zum einen angenommen, dass Experten mehr Reaktionszeitaufgaben richtig lösen (Hypothese 1) und zum anderen auf Bremshinweise, die im Bildmaterial zu den Aufgaben erscheinen, schneller reagieren (Hypothese 2). Beide Hypothesen werden durch die Ergebnisse der zweiten Vorstudie gestützt. In den Reaktionsaufgaben waren die Teilnehmer, die bereits über Fahrerfahrung verfügten (Experten), den Teilnehmern ohne Fahrerfahrung (Novizen) deutlich überlegen. Sie lösten mehr Aufgaben richtig und reagierten tendenziell schneller auf Hinweise, die eine Bremsaktion verlangten. Die Ergebnisse sprechen somit für die Kriteriumsvalidität des Aufgabenformates, da die Aufgaben wirkliche Experten von wirklichen Novizen trennen können. Dies zeigte sich ebenfalls, wenn im Sinne der

Signalentdeckungstheorie (Swets, 1961) der Einfluss möglicher Antworttendenzen auf die Leistungsberechnung kontrolliert wurde.

Die Ergebnisse weisen zudem darauf hin, dass Fahrerfahrung die zuverlässige und schnelle Identifikation relevanter Umweltreize fördert. Für die Identifikation von Gefahrenhinweisen wurde dies bisher bereits in früheren Studien bestätigt (z.B. McKenna & Crick, 1991; McKenna & Crick, 1994a). Die vorliegende Studie zeigt allerdings, dass Verkehrssituationen von Experten auch hinsichtlich weniger salienter Aspekte, wie beispielsweise geltenden Regeln, schneller und zuverlässiger verarbeitet werden.

In Anlehnung an die Ergebnisse der Expertiseforschung, wurde angenommen, dass Experten die Novizen eher übertreffen, je eher das eingesetzte Aufgabenformat den Anforderungen des tatsächlichen Fahrens entspricht. Es wurde zudem vermutet, dass die Aufgaben in ihrer dynamischen Version realitätsnäher sind, da sie zum einen eine situationsnahe Wahrnehmung der bewegten Umwelt ermöglichen und zum anderen eine rechtzeitige Reaktion innerhalb eines kritischen Zeitfensters erfordern, was der tatsächlichen Handlung beim Fahren entspricht. Es wurde daher angenommen, dass Experten die Novizen in der dynamischen Version des Testmaterials deutlicher übertreffen als in der statischen Version des Testmaterials (Hypothese 3). Zudem sollten die Experten die Novizen in der Geschwindigkeit bei der Identifikation der Bremshinweise in der dynamischen Version stärker übertreffen als in der statischen Version des Testmaterials (Hypothese 4). Beide Hypothesen werden durch die Ergebnisse dieser Studie nicht gestützt. Es zeigten sich keine statistisch bedeutsamen Wechselwirkungen zwischen den Faktoren Expertise und Präsentationsform. Beide Testversionen trennten etwa gleich gut zwischen Experten und Novizen. Allerdings zeigte sich, dass die dynamische Version des Testmaterials im Bezug auf die korrekt identifizierten Bremshinweise insgesamt schwieriger war. Das heißt, Experten und Novizen lösten die Aufgaben eher, wenn sie in der statischen Präsentationsform dargeboten wurden. Das dynamische Präsentationsformat scheint demnach insgesamt anspruchsvoller zu sein. Dieser Befund erscheint plausibel, da in der dynamischen Version mehr Reize dargeboten und somit auch verarbeitet werden müssen, was sich auf die Leistung ausgewirkt haben kann. Der Unterschied kann außerdem darauf zurückgeführt werden, dass den Versuchspersonen in der statischen Version unbegrenzt viel Zeit blieb, die Aufgaben zu bearbeiten. Dadurch wurde es den Teilnehmern der statischen Version erleichtert die Aufgaben zu lösen. In der dynamischen Version hingegen gab es nicht die Möglichkeit für die Teilnehmer in unbegrenztem Zeitumfang auf die Videos zu reagieren, denn diese wurden bis

zum Ende vorgespielt und nach Ablauf des Videos konnte nicht mehr reagiert werden. Ein Haupteffekt der Präsentationsform konnte hinsichtlich der Reaktionszeiten nicht festgestellt werden. Die Teilnehmer reagierten demnach in der dynamischen und in der statischen Version etwa gleich schnell auf Bremshinweise. Dass sich die erhöhte Schwierigkeit der dynamischen Aufgaben nicht auf die Reaktionszeiten ausgewirkt hat könnte darin begründet sein, dass die Daten derjenigen Personen, die nicht mehr innerhalb des kritischen Zeitfensters sondern erst später reagiert haben, nicht in die Analyse einbezogen wurden, während es in der statischen Version kein Zeitlimit gab.

Die einzelnen Aufgaben konnten hinsichtlich der Salienz der darin auftretenden Bremshinweise in drei Kategorien eingeteilt werden: Aufgaben ohne Bremshinweis (Distraktoren), Aufgaben mit implizitem Bremshinweis (Vorschriftaufgaben) und Aufgaben mit explizitem Bremshinweis (*Hazard-Perception*-Aufgaben). In Anlehnung an die Ergebnisse von Biermann (2007) wurde angenommen, dass die *Hazard-Perception*-Aufgaben eher richtig gelöst werden als die Vorschriftaufgaben (Hypothese 5) und dass auf die Bremshinweise in den *Hazard-Perception*-Aufgaben schneller reagiert wird als auf die Bremshinweise in den Vorschriftaufgaben (Hypothese 6). Die Ergebnisse der Studie stützen Hypothese 5; Die *Hazard-Perception*-Aufgaben waren insgesamt leichter und wurden häufiger durch die Teilnehmer gelöst als die Vorschriftaufgaben. Es wird daher angenommen, dass die Lösung der Vorschriftaufgaben höhere kognitive Anforderungen an die Teilnehmer stellt, da sie nicht nur die schnelle Wahrnehmung eines relativ deutlichen Gefahrenhinweises verlangt, sondern Wissen um Verkehrsvorschriften und die Anwendung dieses Wissens auf eine komplexe Verkehrssituation verlangt.

Es hat sich außerdem gezeigt, dass die Distraktoren die geringste Schwierigkeit aufwiesen. Diese Tatsache ist mit dem Umstand zu begründen, dass diese Aufgaben bereits mit dem Ziel erstellt wurden, möglichst zweifelsfrei keine Bremshinweise zu enthalten. Die Gewährleistung, dass in einem Verkehrsszenario sicher keine Gefahr vorhanden ist, ist keine triviale Aufgabe. Jeder Verkehrsteilnehmer, der auftaucht, kann prinzipiell relevant für die Verkehrssicherheit werden. Es wurde daher angestrebt, möglichst wenige Verkehrsteilnehmer in den Distraktor-Szenarien auftauchen zu lassen oder diese nur in der Ferne zu zeigen. Dies erklärt die besonders hohe Lösungswahrscheinlichkeit der Distraktoren.

Hypothese 6 konnte durch die Ergebnisse dieser Studie nicht gestützt werden. Es wurde nicht generell schneller auf Bremshinweise in den *Hazard-Perception*-Aufgaben reagiert als in den Vorschriftaufgaben. Wenn Bremshinweise, die sich aufgrund von

Verkehrsvorschriften ergeben, überhaupt erkannt wurden, dann etwa ebenso schnell wie explizite Gefahrenhinweise.

In Anlehnung an die kognitive Flexibilisierungstheorie (Spiro et al., 1988; Spiro et al., 1991) kann davon ausgegangen werden, dass das Wissen von Experten besser vernetzt und schneller abrufbar ist als das Wissen von Novizen. Zudem sollte die Abfrage von bereichsspezifischem Wissen die Aufgaben spezifischer für eine Domäne machen und daher den Unterschied zwischen Experten und Novizen deutlicher hervorheben als bei Aufgaben, die ohne domänenspezifisches Wissen gelöst werden können. Es wurde daher vermutet, dass Experten Novizen deutlicher übertreffen, wenn der Einsatz verkehrsrelevanten Wissens zur Lösung der Aufgaben notwendig ist. Experten sollten demnach die Novizen bei der richtigen Identifikation von Bremshinweisen (Hypothese 7) und bei der Geschwindigkeit der Identifikation von Bremshinweisen (Hypothese 8) innerhalb der Vorschriftaufgaben deutlicher übertreffen als innerhalb der *Hazard-Perception*-Aufgaben. Beide Hypothesen konnten nicht durch die Ergebnisse der Studie gestützt werden. Es konnten keine Wechselwirkungen zwischen dem Faktor Expertise und dem Faktor Aufgabentyp festgestellt werden. Es zeigten sich allerdings signifikante Wechselwirkungen zwischen den Faktoren Aufgabentyp und Präsentationsformat. Dabei stellte sich heraus, dass die Distraktor-Aufgaben und besonders die *Hazard-Perception*-Aufgaben in der dynamischen Testversion schwieriger zu lösen waren als in der statischen Testversion, während das Präsentationsformat bei den Vorschriftaufgaben keinen Einfluss auf die Lösungswahrscheinlichkeit auszuüben schien. Für die Reaktionszeiten konnte festgestellt werden, dass die *Hazard-Perception*-Aufgaben schneller in der dynamischen als in der statischen Version des Testmaterials gelöst wurden, während die Teilnehmer bei den Vorschriftaufgaben in der statischen Version schneller auf relevante Bremshinweise reagierten als in der dynamischen Version. Diese Wechselwirkung wird allerdings durch den dritten Faktor Expertise moderiert. Es zeigte sich dabei, dass die Präsentationsform, in der das Bildmaterial dargeboten wurde, sich eher auf die Leistung der Novizen als auf die Leistung der Experten auswirkt. Besonders bei der Reaktion auf Bremshinweise in Vorschriftaufgaben zeigte sich, dass die Novizen bei einer dynamischen Version später reagierten als in der statischen Version. Dieses Ergebnis könnte dadurch erklärt werden, dass die Vorschriftaufgaben in ihrer statischen Version den Darstellungen auf den Bildern der Aufgaben der derzeitigen theoretischen Fahrerlaubnisprüfung stark ähneln. Bei den meisten Teilnehmern handelte es sich um Fahrschüler, denen dieses Material zum einen bekannt und zum anderen in der Prüfungsvorbereitungsphase besonders vertraut sein

sollte. Diese Vertrautheit mit dem Material ist nicht mehr gegeben, wenn es in seiner dynamischen Version präsentiert wird. Die *Hazard-Perception*-Aufgaben hingegen waren den Novizen in beiden Präsentationsformaten unbekannt. Für die Experten war keines der Präsentationsformate für einen Aufgabentyp bekannter, was die geringeren Leistungsunterschiede plausibel macht.

Die dreifache Interaktion der Faktoren Expertise, Präsentationsform und Aufgabentyp gibt allerdings einen Hinweis drauf, dass der Unterschied zwischen den Experten und den Novizen am deutlichsten ausfällt, wenn die Bedingung vorherrscht, die für die Domäne des Autofahrens am spezifischsten erscheint: Die Probanden bearbeiten in der dynamischen Version des Testmaterials Aufgaben mit implizitem Bremsgrund (Vorschriftaufgaben).

4.2.4.2 Implikationen für das weitere empirische Vorgehen

In der zweiten Vorstudie sollte analog zum Vorgehen in Vorstudie 1 die Validität und Reliabilität eines Aufgabenformates überprüft werden. Eine zufriedenstellende Güte des Aufgabenformates wurde als Voraussetzung zum Einsatz der Aufgaben in einer nachfolgenden Längsschnittuntersuchung erachtet. Zudem sollte entschieden werden, ob die Aufgaben in ihrer statischen oder ihrer dynamischen Präsentationsform in die Längsschnittstudie eingehen sollten.

Insgesamt hat sich das Reaktionszeitaufgabenformat als kriteriumsvalid erwiesen, da die Experten den Novizen darin deutlich überlegen waren. Es konnte damit bestätigt werden, dass sich die Gefahrenwahrnehmungsleistungen mit steigender Fahrerfahrung zu verbessern scheinen (*Hazard-Perception*-Aufgaben). Dieser Effekt der Expertise wurde bereits in früheren Studien gezeigt (z.B. Biermann, 2007; McKenna & Crick, 1991; McKenna & Crick, 1994b; Regan, Triggs & Deery, 1998; Summala & Näätänen, 1988) und konnte in dieser Studie repliziert werden. Zum anderen konnte gezeigt werden, dass sich das Reaktionszeitformat auch zur Erfassung der Fähigkeit zum schnellen Überblicken komplexer Verkehrsszenarien, die keine Hinweise auf direkte Gefahren beinhalten (Vorschriftaufgaben) generell eignet. Das Aufgabenformat bietet daher die Möglichkeit, Wissen über Verhaltensvorschriften im Verkehr handlungsnäher zu überprüfen, als es mit dem derzeit in der theoretischen Fahrerlaubnisprüfung eingesetzten Multiple-Choice-Format möglich ist. Die Fähigkeit der Aufgaben, Experten von Novizen valide zu trennen, lässt vermuten, dass die Aufgaben einen Aspekt von Fahrkompetenz abprüfen und damit auch eine Trennung kompetenter von weniger kompetenten Novizen vermutlich ermöglichen können. Diese Trennung wird mit der theoretischen Fahrerlaubnisprüfung verfolgt, sodass die Aufgaben sich

generell für einen Einsatz in dieser Prüfung eignen. Da die Aufgaben neben ihrer Validität auch eine zufriedenstellende Reliabilität (interne Konsistenz) aufwiesen, wurde entschieden, dieses Aufgabenformat in die Längsschnittuntersuchung einzubeziehen.

Die Betrachtung der Aufgaben, getrennt nach inhaltlichen Aspekten, hatte zum Ergebnis, dass die Distraktoren, in denen kein Bremshinweis enthalten war, eine besonders niedrige Schwierigkeit aufwiesen. Erschweren könnte man die Lösung der Aufgaben lediglich durch den Einbau von Objekten in die Verkehrsszenarien, die nicht eindeutig ungefährlich sind. Genau dies sollte jedoch vermieden werden, da vorsorgliches Bremsen in einem solchen Fall nicht als Fehler gewertet werden sollte. Neben ihrer diagnostischen Funktion dienen die Distraktoren aber auch dazu, die Güte des Aufgabenformates zu sichern. Würden keine Distraktoren einbezogen, wäre den Teilnehmern bewusst, dass bei jedem Bild oder Szenario die Notwendigkeit zur Geschwindigkeitsreduktion gegeben wäre. In der statischen Version hätte dies vermutlich zur Folge, dass die Teilnehmer sofort nach Erscheinen des Bildes reagieren würden. Die Entscheidung ob auf dem Bild ein relevanter Reiz abgebildet ist oder nicht, würde dabei entfallen. In der dynamischen Version wird die Aufgabe durch Existenz von Distraktoren ebenfalls erschwert. Auch hier müsste bei einem Verzicht auf die Distraktoren weniger bewertet werden, ob ein Reiz als relevant oder irrelevant eingeschätzt werden sollte. Daher sollten die Distraktoren in ihrer ursprünglichen Form auch in der Längsschnittstudie zum Einsatz kommen.

Für die Längsschnittstudie sollten zudem einzelne Aufgaben, die eine geringe Lösungswahrscheinlichkeit für die Experten und die Novizen aufweisen, so abgeändert werden, dass eindeutig nur eine Situation auftaucht, die eine Bremsreaktion hervorrufen sollte.

Die Entscheidung, ob das Reaktionszeitformat in seiner statischen oder seiner dynamischen Version in die Längsschnittstudie eingehen sollte, erfolgte unter Berücksichtigung verschiedener Güteaspekte. Es hatte sich gezeigt, dass sich die Kriteriumsvalidität der Aufgaben nicht generell durch die Wahl eines bestimmten Präsentationsformates steigern lässt. Die statische Version ermöglichte die Trennung von Experten und Novizen ebenso gut wie die dynamische Version. Lediglich bei den Vorschriftaufgaben zeigte sich eine deutlichere Trennung zwischen Experten und Novizen, wenn die dynamische Version des Testmaterials vorlag. Dies ist zumindest ein schwacher Hinweis darauf, dass die dynamische Version eine Trennung von Experten und Novizen eher begünstigt. Ausschlaggebend für die Bevorzugung des dynamischen Darstellungsformates

erscheint allerdings die ökologische Validität der Aufgaben, die durch die Dynamik enorm gesteigert werden kann. Dabei werden zwei unterschiedliche Aspekte relevant. Zum einen werden dem Teilnehmer durch die dynamische Darstellung der Verkehrsszenarien Wahrnehmungseindrücke ermöglicht, die eher dem Erleben beim tatsächlichen Fahren entsprechen sollten, als es mit statischen Abbildungen möglich wäre. Zum anderen wird aber auch die Handlungsnähe des Aufgabenformates durch die dynamische Präsentation des Bildmaterials gesteigert. Die Anforderung an die Teilnehmer, genau innerhalb eines relevanten Zeitrahmens zu reagieren, der sich aus der sich stets verändernden Situation ergibt, erscheint eher repräsentativ für die komplexe Fahraufgabe als die Reaktion auf ein Bild. Die dynamische Präsentation und auch das punktgenaue Reagieren auf eine sich stetig verändernde Reizflut sind spezifisch für die Domäne Autofahren. Es wurde daher entschieden, dass die Reaktionszeitaufgaben in ihrer dynamischen Version in die Längsschnittuntersuchung eingehen sollten.

Für die Längsschnittstudie ist wichtig, dass einige Versuchspersonen kritisierten, dass keine Rückmeldung vom Programm gegeben wurde, wenn auf eine Bremsreaktion reagiert wurde. Dabei sollte es sich allerdings nicht um ein Feedback bzgl. der Richtigkeit der Antwort handeln, sondern lediglich um ein Signal dafür, dass die Reaktion des Probanden vom Programm registriert wurde. Dieses Problem sollte für die Längsschnittstudie gelöst werden, indem nach der Reaktion durch den Probanden ein visuelles Zeichen eingeblendet werden sollte.

4.3 Zusammenfassende Diskussion der Vorstudien

Mit dem Ziel der Identifikation von Aufgabenformaten, die erfolgreich zwischen Fahrnovizen und Fahrexperten unterscheiden können, wurden zwei Vorstudien durchgeführt, in denen insgesamt drei Aufgabenarten – Wissens-, Einschätzungs- und Reaktionszeitaufgaben – in Experten-Novizen-Vergleichen hinsichtlich ihrer Kriteriumsvalidität getestet wurden. Zudem wurde überprüft, inwieweit sich das Präsentationsformat (statisch vs. dynamisch), in dem das Bildmaterial, das jede einzelne Aufgabe begleitet, dargeboten wurde, auf die Leistung in den verschiedenen Aufgabenarten auswirkte.

Es hat sich gezeigt, dass lediglich das Reaktionszeitformat zuverlässig zwischen Fahrexperten und Novizen zu trennen vermag. Deutlich überlegen waren die Experten den Novizen darin, relevante Bremshinweise in dargestellten Verkehrsszenarien zu identifizieren (Lösungswahrscheinlichkeit) und ebenfalls in der Geschwindigkeit, in der sie auf die

Bremshinweise reagierten (Reaktionszeit). Die Kriteriumsvalidität dieses Aufgabenformates kann daher als gesichert betrachtet werden. In den Einschätzungs- und Wissensaufgaben übertrafen die Experten die Novizen nicht in statistisch bedeutsamem Ausmaß. Die Einschätzungsaufgaben wiesen neben der fehlenden Validität auch eine mangelnde Reliabilität auf. Es kann daher davon ausgegangen werden, dass mit diesen Aufgaben weder der Einsatz von Faustregeln, noch irgendein anderer Aspekt von Fahrexpertise zuverlässig gemessen werden kann. Dies wurde zum Anlass genommen, den Einsatz dieses Aufgabenformates nicht weiter zu verfolgen. Die Wissensaufgaben hingegen waren reliabel, ermöglichten allerdings ebenfalls keine valide Trennung von Experten und Novizen. Die Tatsache, dass die Experten den Novizen in den Wissensaufgaben nicht deutlich überlegen waren, könnte durch den Umstand erklärt werden, dass die Aufgaben zum Prüfungstoff der theoretischen Fahrerlaubnisprüfung gehören, auf den sich wahrscheinlich diejenigen Teilnehmer, die bei ihrem Fahrerlaubniserwerb schon weit vorangeschritten waren, bereits vorbereitet hatten. Die Novizen waren demnach nicht mehr als absolute Novizen zu bezeichnen. Aus den dargestellten Ergebnissen lässt sich für eine geplante Längsschnittstudie, die das Lernen der Novizen während ihrer Fahrschulausbildung und kurz danach abbilden soll, klar erkennbar ableiten, dass das Reaktionszeitformat eingesetzt werden sollte, während auf die Einschätzungsaufgaben verzichtet werden sollte. Die Wissensaufgaben sollten ebenfalls in die Längsschnittstudie eingehen, da vermutet werden kann, dass die Experten den Novizen zumindest zum Beginn von deren Ausbildung überlegen sind. Daraus ergibt sich die Voraussetzung für die Längsschnittstudie, dass die Novizen sich zum ersten Messzeitpunkt ganz am Anfang ihrer Fahrschulausbildung befinden sollten.

Neben der generellen Identifikation von kriteriumsvaliden Aufgabenformaten wurde bei den Vorstudien das Ziel verfolgt, zu überprüfen, in welchem Präsentationsformat – statisch oder dynamisch – das Bildmaterial zu den Aufgaben dargeboten werden sollte um eine optimale Trennung zwischen Experten und Novizen zu gewährleisten. Allerdings hat sich gezeigt, dass die Aufgaben in keinem der beiden Präsentationsformate generell besser zwischen Novizen und Experten zu trennen vermögen. Eine Verbesserung der Kriteriumsvalidität durch den Einsatz eines bestimmten Präsentationsformates, die eigentlich als Entscheidungshilfe für die Wahl des Präsentationsformates für die Aufgaben der Längsschnittstudie dienen sollte, konnte somit für keine Aufgabenart festgestellt werden. Die Entscheidung für ein Präsentationsformat musste daher nach anderen Kriterien erfolgen. Es wurde festgestellt, dass das Präsentationsformat in Abhängigkeit von der jeweiligen

Aufgabenart einen Einfluss auf die Leistungen zu haben scheint. Es hat sich gezeigt, dass die Aufgaben zum deklarativen Wissen eher von Novizen gelöst wurden, wenn sie dynamisch präsentiert wurden. Dies weist darauf hin, dass das Instruktionsverständnis gesteigert werden konnte, was für den Einsatz der Aufgaben in ihrer dynamischen Form spricht. Durch das gesteigerte Instruktionsverständnis sollte der Messfehler der Skala gesenkt werden. Dies führt vermutlich zu einer Erhöhung der Reliabilität und Validität der Aufgaben, da der Anteil wahrer Varianz an der Gesamtvarianz der Messergebnisse und damit ihre Genauigkeit steigt und zum anderen die Gefahr eines Testbias reduziert wird. Zudem kann die inhaltliche Validität der Aufgaben durch die dynamische Präsentation des Bildmaterials gesteigert werden, da die durch dieses Präsentationsformat bereitgestellten visuellen Eindrücke eher denen beim realen Fahren entsprechen, als es bei statischen Abbildungen der Fall sein sollte. Basierend auf den dargestellten Vorteilen der dynamischen Präsentation, wurde die Entscheidung getroffen, die Wissensaufgaben in der Längsschnittstudie mit animierten Verkehrsszenarien darzubieten.

Im Falle der Reaktionszeitaufgaben hat sich nicht gezeigt, dass generell mit einer Verbesserung der Kriteriumsvalidität durch den Einsatz eines bestimmten Präsentationsformates zu rechnen ist. Die Tatsache, dass in Abhängigkeit vom Inhalt einige Aufgaben (Vorschriftaufgaben) in ihrer dynamischen Version zwischen Experten und Novizen valide trennen konnten und in der statischen Version nicht, gibt einen Hinweis darauf, dass die dynamische Version prinzipiell das Potenzial zur Validitätssteigerung birgt. Zudem wird durch die dynamische Darbietung des Bildmaterials die ökologische Validität der Aufgaben erhöht. Die Anforderungen an die Testteilnehmer sind im dynamischen Format höher und repräsentativer für die reale Fahraufgabe. Die durch den Einsatz von animierten Verkehrsszenarien erzielte Validitätssteigerung führte zu der Entscheidung, auch die Reaktionszeitaufgaben in der Längsschnittstudie in ihrer dynamischen Version darzubieten.

In den beiden Vorstudien wurden als Personenmerkmale räumliche Fähigkeiten (Visualisierungsfähigkeit) und die Sensationssuche erhoben, um den Einfluss dieser Merkmale auf die Leistung in den unterschiedlichen Aufgabenformaten zu erfassen. Für die Wissensaufgaben hat sich gezeigt, dass die Fähigkeit zur räumlichen Visualisierung insgesamt in einem schwachen positiven Zusammenhang mit der Leistung in den Testaufgaben stand. Der Einfluss der Visualisierungsfähigkeit als Kovariate war statistisch nicht bedeutsam. Stärkeren Einfluss könnte die Visualisierungsfähigkeit jedoch bei absoluten Novizen haben. Bei den Novizen, die an der Längsschnittstudie teilnehmen, sollte es sich zum ersten

Messzeitpunkt um absolute Novizen ohne fahr- und verkehrsspezifisches Wissen handeln. Daher soll die Visualisierungsfähigkeit in der Längsschnittstudie ebenfalls gemessen werden. Der Zusammenhang zwischen der Leistung in den Reaktionszeitaufgaben und der Visualisierungsfähigkeit wurde in den Vorstudien nicht überprüft. Dies sollte daher in der Längsschnittstudie erfolgen. Der in der ersten Vorstudie eingesetzte Paper Folding Test zur Erfassung der räumlichen Visualisierungsfähigkeit sollte jedoch ergänzt werden durch einen Test, der die räumlichen Fähigkeiten einer Person unter Einbezug einer Zeitkomponente misst. Dies entspricht eher den Anforderungen bei den Reaktionszeitaufgaben sowie beim realen Fahren selbst. Daher sind deutlichere Zusammenhänge zu erwarten.

Zwischen *Sensation Seeking* – einem Persönlichkeitsmerkmal, das in der zweiten Vorstudie erfasst wurde – und der Leistung in den Reaktionszeitaufgaben zeigten sich keine statistisch bedeutsamen Zusammenhänge. Der Grund dafür könnte die fehlende Nähe der ausgewählten Situationen, die in den Items der *Sensation-Seeking*-Skala aufgeführt sind, zum Autofahren sein. In der Längsschnittstudie soll daher eine Skala zur Erfassung der Sensationssuche eingesetzt werden, deren Items sich direkt auf das Autofahren beziehen.

5 Längsschnittstudie

5.1 Ziele und Fragestellungen

In den Vorstudien 1 und 2 wurden verschiedene Aufgabenformate auf ihre Kriteriumsvalidität hin geprüft. Damit konnte ermittelt werden, welche Aufgabenformate Aspekte von Fahrexpertise messen können und daher möglicherweise auch zur Abbildung des Erwerbs von Fahrexpertise dienen können. Zur Untersuchung des Expertiseerwerbs erscheinen Längsschnittstudien dienlich, in denen erfasst wird, inwieweit sich Probanden über die Zeit in domänenspezifischen Aufgaben verbessern. Im Folgenden werden zunächst die Ziele dargestellt, die mit der Durchführung einer Längsschnittstudie zum Erwerb von Expertise in der Domäne Autofahren und Verkehr, verfolgt wurden. Anschließend werden die Hypothesen dargelegt.

5.1.1 Ziele

In der Längsschnittstudie sollte überprüft werden, wie sich die Leistung von Fahrexperten und Fahrnovizen in ausgewählten Aufgabenformaten über die Zeit entwickelt. Von besonderem Interesse war dabei die Entwicklung der unerfahrenen Fahrer, die vom Beginn ihrer Fahrausbildung bis einige Zeit nach dem Erwerb der Fahrerlaubnis begleitet wurden. Die Längsschnittstudie sollte zudem Einblicke in die beginnende Entwicklung verschiedener Aspekte der Fahrexpertise gewähren.

In den Vorstudien hat sich gezeigt, dass sich das Präsentationsformat (statisch vs. dynamisch) bei keinem der Aufgabenformate auf die Leistung der Teilnehmer in Abhängigkeit von ihrer Expertise ausgewirkt hat. Auch wenn damit keine generelle Steigerung der Kriteriumsvalidität der Aufgaben durch den Einsatz von dynamisch dargebotenen Verkehrsszenarien gezeigt werden konnte, wurden andere Vorteile einer dynamischen Vorgabe des Testmaterials identifiziert (vgl. Kapitel 4.3). In der Längsschnittstudie wurde daher darauf verzichtet den Faktor Präsentationsformat zu variieren. Alle Aufgabenformate mit Bildmaterial waren durch dynamische Darstellungen von Verkehrsszenarien angereichert.

Bei der Längsschnittstudie handelte es sich um einen Experten-Novizen-Vergleich mit drei Messzeitpunkten. Wie in den Vorstudien wurden als Experten Personen bezeichnet, die bereits seit zwei Jahren oder länger eine Fahrerlaubnis der Klasse B besaßen. Die Novizen waren Fahrschüler, die sich am ersten Messzeitpunkt ganz zu Beginn ihrer

Fahrschulausbildung befinden sollten. Bis zur ersten Erhebung durften maximal vier praktische Fahrstunden absolviert werden und die Theorieprüfung noch nicht abgelegt sein. Die Messungen fanden in einem Abstand von etwa drei bis vier Monaten statt. Daher sollten sich die Novizen beim zweiten Messzeitpunkt kurz vor oder kurz nach dem Erwerb ihrer Fahrerlaubnis befinden. Am dritten Messzeitpunkt sollten sie bereits etwa drei Monate lang selbstständig gefahren sein und damit etwas praktische Fahrerfahrung gesammelt haben. Die drei Messzeitpunkte stellen kritische Punkte während der Ausbildung von Fahrexpertise dar: Während zum ersten Messzeitpunkt weder eine große theoretische Wissensbasis noch praktisches Können aufgebaut sein sollte, wurde bis zum zweiten Messzeitpunkt bereits eine breite Wissensbasis aufgebaut und etwas praktische Fahrerfahrung in wenigen Stunden erworben. Bis zum dritten Messzeitpunkt sollten die Fahrschüler durch selbstständiges Fahren besondere Fortschritte in Aspekten der Fahrkompetenz gemacht haben, die in deutlichem Zusammenhang mit dem tatsächlichen Verhalten im Verkehr stehen. Fahrexperten, die bereits über ein hohes Maß an fahr- und verkehrsspezifischer Expertise und Erfahrung verfügen, sollten sich hingegen über die Zeit hinweg kaum in ihrer Leistung verbessern. Durch die Längsschnittstudie sollte ermittelt werden, ob sich der theoretisch angenommene beginnende Aufbau von Expertise bei den Fahrnovizen in einer Leistungssteigerung über die Zeit äußert. Dies spräche dafür, dass die ausgewählten Aufgaben sensibel für Veränderungen in verschiedenen Aspekten der Fahrexpertise sind und sich damit zur Messung dieser Veränderungen eignen. Insbesondere sollte dabei zwischen verschiedenen Aufgabenarten, die konzipiert wurden um unterschiedliche Aspekte der Fahrexpertise zu messen (deklaratives Faktenwissen, regelbasiertes Wissen, angemessene Reaktion auf Verkehrsgegebenheiten), unterschieden werden.

5.1.2 Hypothesen

5.1.2.1 Hypothesen zu den Experten-Novizen-Vergleichen

Die Novizen sollten während der Zeit der Studie ihre Fahrausbildung in der Fahrschule absolvieren. Da sie während dessen durch den Theorieunterricht fahrerspezifisches Wissen und durch das Fahren im Realverkehr Wissen und Fertigkeiten erwerben sollten, werden Hypothesen bzgl. des Unterschiedes zwischen Experten und Novizen getrennt für die drei Messzeitpunkte und für die verschiedenen Aufgabenformate formuliert.

Messzeitpunkt 1

Zum Messzeitpunkt 1 sollten die Novizen erst seit kurzem die Fahrschule besuchen und daher noch kaum Verkehrswissen oder Faustregeln erworben haben. Die Experten, die bereits über zwei Jahre Fahrerfahrung gesammelt haben, sollten hingegen mehr Fakten und Regeln bezüglich des Fahrens kennen. Ein Merkmal von Experten ist, dass sie in ihrer Domäne über besonders viel Wissen verfügen (vgl. z.B. Gruber & Mandl, 1995). Da die Novizen sich bedingt durch ihre fehlende Ausbildung auf einem sehr niedrigen Kenntnisstand und die Experten sich dem gegenüber auf einem hohen Kenntnisstand befinden sollten, ist gemäß den Befunden der Expertiseforschung in anderen Domänen davon auszugehen, dass die Experten die Novizen deutlich übertreffen. In Vorstudie 1 konnte diese Annahme nicht bestätigt werden: Experten und Novizen erbrachten ähnliche Leistungen in den Bereichen Wissen und Kenntnis von Faustregeln. Dieses unerwartete Ergebnis wurde allerdings darauf zurückgeführt, dass es sich bei den Novizen nicht mehr um absolute Novizen handelte. Sie hatten zum Teil bereits am theoretischen und praktischen Fahrschulunterricht teilgenommen, was in der Längsschnittstudie ausgeschlossen wurde, da sich die Novizen bei Beginn der Studie erst am Anfang ihrer Fahrschulausbildung befanden.

Hypothese 1: Die Experten lösen zu Messzeitpunkt 1 mehr Wissensaufgaben richtig als die Novizen.

Hypothese 2: Die Experten lösen zu Messzeitpunkt 1 mehr Regelaufgaben richtig als die Novizen.

In zahlreichen Studien konnte belegt werden, dass erfahrene Fahrer den unerfahrenen Fahrern in sogenannten *Hazard-Perception*-Aufgaben, bei denen das Absuchen der Umwelt nach sicherheitsrelevanten Reizen verlangt wird, überlegen sind (vgl. Abschnitt 2.3.3.1). Das heißt, Gefahren werden von erfahrenen Fahrern eher als gefährlich eingeschätzt und schneller erkannt. Die Ergebnisse der Vorstudie 2 belegen dies ebenfalls und weisen zudem darauf hin, dass neben offensichtlichen Gefahren besonders auch andere Gegebenheiten (z.B. Vorfahrtsregeln), die eine Geschwindigkeitsreduktion verlangen, von Experten zuverlässiger und schneller erkannt werden als von Novizen. Diese Befunde sind konsistent mit den Befunden der Expertiseforschung in anderen Domänen, wonach Experten schneller und flexibler spezifische Probleme lösen (vgl. z.B. Glaser & Chi, 1988; Gruber & Mandl, 1995).

Daher ist anzunehmen, dass die Experten die Novizen im Reaktionszeitaufgabenformat zu Messzeitpunkt 1 übertreffen.

Hypothese 3: Die Experten lösen zu Messzeitpunkt 1 mehr Reaktionszeitaufgaben richtig als die Novizen.

Hypothese 4: Die Experten reagieren zu Messzeitpunkt 1 in den Reaktionszeitaufgaben schneller als die Novizen.

Messzeitpunkte 2 und 3

Um Voraussagen bzgl. der Leistungen der erfahrenen und unerfahrenen Fahrer zu den beiden späteren Messzeitpunkten 2 und 3 vornehmen zu können, ist es erforderlich einzubeziehen, dass die Novizen im Gegensatz zu den Experten in den Zeiträumen zwischen den Messzeitpunkten Anstrengungen unternehmen, ihre Fahr- und Verkehrsexpertise zu verbessern. Sie besuchen den theoretischen Fahrschulunterricht, lernen für die theoretische Prüfung und absolvieren praktische Fahrstunden, in denen sie Feedback von ihren Fahrlehrern bezüglich ihres Fahrverhaltens erhalten. Problematische Situationen werden dabei wenn möglich häufiger aufgesucht und gezielt geübt. Zum Messzeitpunkt 2 sollten alle Novizen mindestens die theoretische Fahrerlaubnisprüfung absolviert haben. Relevantes deklaratives Verkehrswissen wurde innerhalb der Theoriestunden erworben und in den meisten Fällen durch eigenständiges Bearbeiten von Übungsbögen bzw. Trainingsprogrammen gefestigt. Zudem sollten einige Novizen auch schon eine praktische Fahrerlaubnisprüfung absolviert haben, was ein gewisses Ausmaß an praktischer Fahrerfahrung (mindestens die 12 gesetzlich vorgeschriebenen Sonderfahrten à 45 Minuten) voraussetzt. Zum Messzeitpunkt 3 sollten die meisten Novizen ihre Fahrausbildung komplett absolviert haben und schon kurze Zeit selbstständig gefahren sein.

Mit der theoretischen Fahrausbildung wird das Ziel verfolgt, relevante Inhalte bzgl. des Verkehrswissens zu vermitteln. Das Wissen der Novizen um Gefahren im Verkehr und Faustregeln sollte demnach durch die Ausbildung dem Wissen der Experten entsprechen. Hier sind keine Unterschiede zu erwarten und folglich keine Unterschiedshypothesen zu treffen. Die Fähigkeit zum Transfer und zum flexiblen Einsatz von erlerntem Wissen und Regeln auf andere Situationen bildet sich erst durch Erfahrung aus (vgl. z.B. Dreyfus, S. E. & Dreyfus, 1980; Spiro et al., 1988). Die Novizen sollten ihr Wissen aufgrund ihrer geringen Erfahrung noch nicht in genügend Verkehrssituationen eingesetzt haben, um es flexibilisiert zu haben.

Daher sollten Experten den Novizen in denjenigen Aufgabenformaten, die handlungsnah einen Transfer ihres Wissens verlangen, auch nach Abschluss der theoretischen und praktischen Fahrausbildung überlegen sein.

Hypothese 5: Die Experten lösen zu Messzeitpunkt 2 und 3 mehr Reaktionszeitaufgaben richtig als die Novizen.

Hypothese 6: Die Experten reagieren in den Reaktionszeitaufgaben zu Messzeitpunkt 2 und 3 schneller als die Novizen.

5.1.2.2 Hypothesen zur Leistungsentwicklung über die Zeit

Innerhalb der Längsschnittuntersuchung wird die Leistung der Novizen und Experten über die Zeit betrachtet. Über die Leistungsentwicklung lassen sich verschiedene Annahmen für die unterschiedlichen Teilaspekte der Fahrexpertise und damit für die unterschiedlichen Aufgabentypen treffen. Eine Visualisierung der Hypothesen zur Leistungsentwicklung über die Zeit findet sich in Kapitel 3.2 (Abbildung 3.2).

Das Wissen, das zur Lösung der Wissens- und Regelaufgaben benötigt wird, wird den Fahrschülern während ihrer Fahrausbildung in der Fahrschule nähergebracht und in der Prüfung abgeprüft. Während die Novizen zu Messzeitpunkt 1 – also zu Beginn ihrer Fahrausbildung – kaum Verkehrswissen haben sollten, sollten sie nach ihrer Fahrausbildung (Messzeitpunkt 2 und 3) alles relevante Wissen erworben haben. Die Novizen bereiten sich im Sinne der *deliberate-practice*-Theorie zum Expertiseerwerb in optimaler Weise auf die Prüfung vor (vgl. Kapitel 2.1.3.2): Sie werden im Theorieunterricht regelmäßig instruiert, zwischen den Theoriesitzungen erfolgt freies Üben mit dem Übungsmaterial und direktes Feedback erhalten sie bei der eigenständigen Auswertung ihrer Leistung in den Übungsbögen. Das Wissen der Novizen sollte somit über die Zeit linear ansteigen. Experten sollten über die drei Messzeitpunkte hinweg hingegen ein eher stabiles Verkehrswissen aufweisen, da sie in der Regel bereits über die relevanten Kenntnisse verfügen und sich daher über die Zeit in geringerem Maße verbessern als die Novizen.

Hypothese 7: Die Novizen steigern ihre Leistung in den Wissensaufgaben deutlicher als die Experten.

Hypothese 8: Die Novizen steigern ihre Leistung in den Regelaufgaben deutlicher als die Experten.

Zur Lösung der Reaktionszeitaufgaben werden zum einen Verkehrswissen und -regeln benötigt, zum anderen auch Erfahrung, die es erlaubt, dieses Wissen zunehmend auf unterschiedliche Situationen zu übertragen und angemessen anzuwenden. Novizen sammeln während ihrer Fahrausbildung und in der Zeit danach Wissen und Erfahrung. Sie sollten sich demnach über die drei Messzeitpunkte in den Aufgaben steigern. Die Experten hingegen sollten sich weniger in den Aufgaben steigern, da sie sich bereits auf einem hohen Leistungsniveau befinden.

Hypothese 9: Die Novizen steigern ihre Leistung in den Reaktionszeitaufgaben bzgl. der richtig identifizierten Bremshinweise über die Zeit deutlicher als die Experten.

Hypothese 10: Die Novizen steigern ihre Leistung in den Reaktionszeitaufgaben bzgl. der Reaktionszeiten über die Zeit deutlicher als die Experten.

In Vorstudie 2 hat sich gezeigt, dass die drei verschiedenen Aufgabentypen im Reaktionstest unterschiedlich schwierig sind und mit den Faktoren Expertise und Präsentationsform interagieren. Bisher wurde noch nicht untersucht, wie sich die Leistung in Abhängigkeit vom Aufgabentyp über die Zeit bei Experten und Novizen entwickelt.

Forschungsfrage 1: Wie entwickelt sich die Leistung der Experten und Novizen in den verschiedenen Aufgabentypen im Reaktionszeitaufgabenformat über die Zeit?

Räumliche Fähigkeiten spielen eine moderierende Rolle beim Lernen mit dynamischem Lernmaterial (Höffler, 2010). Welche Rolle diese Fähigkeiten bei der Bearbeitung von Prüfungsmaterial in Verbindung mit der Entwicklung von Expertise spielt, ist derzeit noch nicht geklärt.

Forschungsfrage 2: Liegen Zusammenhänge zwischen der Leistung der Experten und Novizen in den unterschiedlichen Aufgabenformaten und den räumlichen Fähigkeiten vor?

Forschungsfrage 3: Ändert sich die Höhe des Zusammenhangs zwischen den räumlichen Fähigkeiten und der Leistung in verkehrsspezifischen Aufgaben, wenn die Fahr- und Verkehrsexpertise zunimmt?

5.2 Methode

5.2.1 Experimentelles Design und Stichprobe

Bei der Studie handelte es sich um ein 2x3-faktorielles Längsschnittdesign mit Messwiederholung auf einem Faktor. Bei der ersten unabhängigen Variablen handelte es sich um den Faktor Expertise mit den beiden Ausprägungen Experte (Fahrerlaubnis seit mindestens zwei Jahren) und Novize (keine Fahrerlaubnis am ersten Messzeitpunkt). Die zweite unabhängige Variable stellte der Messwiederholungsfaktor Zeit dar. Der Faktor Zeit lag dreifach gestuft vor, wobei die einzelnen Stufen Messzeitpunkte im Abstand von etwa drei bis vier Monaten repräsentieren. Abhängige Variablen waren die Leistungen in den Aufgaben zum Verkehrswissen, den Regelaufgaben und den Reaktionszeitaufgaben.

Die Berechnung der optimalen Stichprobengröße erfolgte nach Faul, Erdfelder, Lang und Buchner (2007). Dazu wurde das Signifikanzniveau auf 5 % festgelegt. Die Schätzung der Effektgröße orientiert sich an den Ergebnissen der Experten-Novizen-Vergleiche der beiden Vorstudien. Für die unterschiedlichen Aufgabenformate wurden niedrige bis mittelhohe Effektstärken ermittelt. Wie groß der Effekt der Zeit eingeschätzt werden kann, ist bisher unbekannt. Daher wurde ein mittlerer Effekt festgelegt, der mit einer Wahrscheinlichkeit von mindestens 90 % gefunden werden sollte. Aus diesen Überlegungen resultiert für diese Studie ein optimaler Gesamtstichprobenumfang von $N = 70$. Bei Längsschnittstudien ist in der Regel mit hohen Ausfallquoten zu rechnen. Um eine ausreichende Teststärke zu gewährleisten, wurde entschieden, eine mindestens doppelt so große Anzahl an Versuchspersonen zu rekrutieren. Die Stichprobe der Längsschnittstudie bestand daher aus insgesamt 154 Personen. Davon nahmen 145 Personen zu Messzeitpunkt 1 zum ersten Mal teil. Von diesen waren 83 Personen Fahrschüler (Novizen), und 62 Personen, die bereits seit zwei Jahren oder länger im Besitz einer Fahrerlaubnis waren (Experten). Neun Personen (allesamt Fahrschüler) nahmen erst am 2. Messzeitpunkt zum ersten Mal teil. Es war notwendig, diese neun Personen zusätzlich in die Studie einzubeziehen, da sich zu Beginn der zweiten Erhebungsphase bereits abzeichnete, dass ein großer Teil der Fahrschüler nicht zur weiteren Teilnahme an der Längsschnittstudie bereit war. Am zweiten Messzeitpunkt haben insgesamt 126 Personen teilgenommen. Zu Messzeitpunkt 3 nahmen 120 Personen teil. Keine der Personen nahm zu Messzeitpunkt 3 zum ersten Mal teil, aber sechs Personen nahmen wieder teil, nachdem sie Messzeitpunkt 2 verpasst hatten. In Tabelle 5.1 sind die Stichprobenzahlen pro Messzeitpunkt dargestellt.

Tabelle 5.1

Stichprobe getrennt nach Geschlecht und Expertise zu den einzelnen Messzeitpunkten (N = 154)

		MZP1 n = 145	MZP2 n = 126 ^{a)}	MZP3 n = 120 ^{b)}
Expertise	Geschlecht			
	Weiblich	53	41	39
	Novizen			
	Männlich	30	22	21
	Gesamt	83	63	60
Experten	Weiblich	36	32	31
	Männlich	26	22	23
	Gesamt	62	54	54

Anmerkungen. ^{a)} Das Geschlecht war nur von 117 Personen bekannt, da neun Novizen ihr Geschlecht nicht angaben. ^{b)} Das Geschlecht war nur von 114 Personen bekannt, da sechs Novizen ihr Geschlecht nicht angaben.

Insgesamt haben 100 Personen (50 % Novizen) an jedem der drei Messzeitpunkte der Längsschnittstudie teilgenommen.

Es lagen drei Parallelversionen des gesamten Testmaterials vor. An jedem Messzeitpunkt bearbeiteten die Personen eine andere Paralleltestversion. Es wurde eine Gleichverteilung der drei Paralleltestformen zu jedem der drei Messzeitpunkte angestrebt. Aus Tabelle 5.2 wird ersichtlich, dass dies zu allen 3 Messzeitpunkten nicht exakt, sondern nur annähernd gelungen ist.

Tabelle 5.2

Häufigkeitsverteilungen (%) der Experten und Novizen auf die unterschiedlichen Paralleltestversionen an den unterschiedlichen Messzeitpunkten

Expertise	Paralleltest	MZP1	MZP2	MZP3	Gesamt
Novizen	A	24 (17)	31 (25)	18 (15)	73 (19)
	B	27 (19)	15 (12)	28 (23)	70 (18)
	C	32 (22)	26 (21)	19 (16)	77 (21)
Experten	A	22 (15)	17 (14)	18 (15)	57 (14)
	B	19 (13)	24 (19)	14 (12)	57 (14)
	C	21 (14)	13 (10)	23 (19)	57 (14)
Gesamt	A, B, C	145 (100)	126 (100)	120 (100)	391 ^{a)} (100)

Anmerkung. ^{a)} Anzahl der Messungen insgesamt

Vor den jeweiligen Messungen wurden die Namen der Teilnehmer zufällig gleichmäßig den Paralleltestversionen zugeordnet. Die Ursache dafür, dass dennoch nicht alle Paralleltestformen an allen Messzeitpunkten in gleicher Häufigkeit eingesetzt wurden, liegt in dem Umstand begründet, dass zu jedem Messzeitpunkt potenzielle Teilnehmer, die sich zuvor gemeldet hatten, nicht erschienen.

Die Versuchspersonen waren im Durchschnitt 21.17 Jahre alt ($SD = 6.5$), wobei die Novizen zwischen 16 und 26 Jahren ($M = 17.57$, $SD = 2.14$) und die Experten zwischen 19 und 53 Jahren ($M = 25.95$, $SD = 7.3$) alt waren. Somit waren die Experten signifikant älter als die Novizen ($T(147) = 9.85$; $p \leq .001$; $d = 1.54$).

Die nun folgenden demographischen Stichprobendaten bzgl. Einwohnerzahl des Wohnortes, höchstem Schulabschluss und beruflicher Situation wurden zum ersten Messzeitpunkt erhoben und beziehen sich daher nur auf die 145 Personen, die an diesem Messzeitpunkt zum ersten Mal an der Studie teilgenommen hatten. Die meisten Teilnehmer wohnten in kleineren Dörfern (37 %). Aus Tabelle 5.3 wird ersichtlich, dass auch viele Personen in Kleinstädten bis 50 000 Einwohner wohnten. Bis auf eine Person handelte es sich bei allen Personen, die in Großstädten ab 100 000 Einwohnern wohnten, um Experten.

Tabelle 5.3

Einwohnerzahlen der Wohnorte der Versuchspersonen (%)

Einwohnerzahl Wohnort	Novizen n = 83	Experten n = 62	Gesamt N = 145
< 5000	34 (41)	20 (32)	54 (37)
5000 - 10 000	16 (19)	11 (18)	28 (19)
10 000 – 50 000	27 (33)	15 (24)	43 (30)
50 000 – 100 000	5 (6)	1 (2)	6 (4)
100 000 – 200 000	0 (0)	13 (21)	13 (9)
> 200 000	1 (1)	2 (3)	3 (2)

Der größte Anteil der Personen insgesamt gab als höchsten Bildungsabschluss die Mittlere Reife an. Aus Tabelle 5.4 wird allerdings ersichtlich, dass der Großteil der Experten bereits die allgemeine Hochschulreife erworben hatte. Ausschließlich unter den Novizen befanden sich Personen, deren höchster Bildungsabschluss der Hauptschulabschluss war, bzw. die bisher keinen Schulabschluss erworben hatten.

Tabelle 5.4
Höchster Bildungsabschluss der Versuchspersonen (%)

Höchster Bildungsabschluss	Novizen n = 83	Experten n = 62	Gesamt N = 145
(noch) kein Abschluss	3 (4)	0 (0)	3 (2)
Hauptschulabschluss	30 (36)	0 (0)	30 (20)
Mittlere Reife	47 (57)	11 (18)	58 (40)
Fachabitur	2 (2)	8 (13)	10 (7)
Abitur	1 (1)	39 (63)	40 (28)
Hochschul- oder Fachhochschulabschluss	0 (0)	4 (6)	4 (3)

Die Diskrepanz zwischen Novizen und Experten beim Bildungsabschluss ist darauf zurückzuführen, dass die Novizen ein geringeres Lebensalter als die Experten aufwiesen und es sich bei einem großen Teil noch um Personen in der Schulausbildung handelte, die ihren höchsten Bildungsabschluss noch nicht erreicht hatten.

Diese Annahme wird durch die Tatsache gestützt, dass es sich, wie aus Tabelle 5.5 ersichtlich, bei den meisten Novizen um Schüler handelte.

Tabelle 5.5
Berufliche Situation der Versuchspersonen (%)

Berufliche Situation	Novizen n = 83	Experten n = 62	Gesamt N = 145
Schüler/in	68 (82)	5 (8)	73 (50)
Zurzeit in der Lehre	7 (8)	11 (18)	18 (12)
Lehre gerade abgeschlossen	1 (1)	0 (0)	1 (1)
In sonstiger Berufsausbildung	1 (1)	4 (6.5)	5 (3)
Voll berufstätig	3 (4)	5 (8)	8 (6)
arbeitslos	3 (4)	4 (6.5)	7 (5)
Student/in	0 (0)	33 (53)	33 (23)

Bei den Experten handelte es sich beim größten Teil um Studenten der Universität des Saarlandes und der Fachhochschule Zweibrücken.

5.2.2 Versuchsmaterial

5.2.2.1 Verkehrsspezifische Leistungsmessung (Paralleltests)

Innerhalb der Längsschnittstudie sollte überprüft werden, wie sich die Leistung der Versuchsteilnehmer in den entwickelten Aufgabenformaten über die Zeit verändert. Die in den Vorstudien getesteten und dargestellten Wissens- und Reaktionszeitaufgaben wurden zur verkehrsspezifischen Leistungsmessung in die Längsschnittstudie einbezogen (siehe 4.1.2 für die Beschreibung der Wissensaufgaben und 4.2.2 für die Beschreibung der Reaktionszeitaufgaben). Einige Szenarien, in denen die Versuchspersonen in den Vorstudien durch das Auftauchen ambivalenter Reize dazu veranlasst wurden, vor dem Auftauchen des kritischen Reizes zu reagieren, wurden leicht verändert. Es wurde dabei darauf geachtet, dass keine möglicherweise zu einer Reaktion führenden Reize vor dem kritischen mehr auftauchten.

Um Erinnerungseffekte zu vermeiden, die vor allem für die Aufgaben im Reaktionszeitformat zu erwarten waren, wurden zu jeder Aufgabe zusätzlich zwei parallele Versionen erstellt. Dabei wurden ausschließlich die eingesetzten Animationen verändert. Aufgabenstamm und Antwortalternativen der Aufgaben im Multiple-Choice-Antwortformat waren hingegen in den jeweiligen Parallelaufgaben identisch. Insgesamt beschränkten sich die Unterschiede zwischen den drei Versionen der jeweiligen Aufgabe auf die Variation von Umweltmerkmalen innerhalb der Animationen der Verkehrsszenarien. Unterschiede wurden hinsichtlich der Art, Anzahl und Platzierung der dargestellten Häuser oder Bäume erzeugt. Zudem wurden die Fahrzeugtypen und -farben variiert und Unterschiede bei den im jeweiligen Szenario auftauchenden Personen geschaffen (Geschlecht, Kleidung, Alter, Anzahl, Geschwindigkeit). Seltener wurden Witterung und Tageszeit (Morgen, Mittag, Abend) verändert. Die Komplexität sollte allerdings bei allen drei Versionen des jeweiligen Szenarios vergleichbar sein, da darauf geachtet wurde, dass die Gesamtanzahl der Objekte in den Animationen aller Parallelaufgaben einheitlich war.

Die Aufgaben wurden zu drei Paralleltestversionen (A, B und C) zusammengestellt. Dabei wurde per Zufall entschieden, welche der drei Versionen eines Szenarios dem jeweiligen Paralleltest zugeordnet wurde.

Die Faustregeln für Sicherheitsabstand, Anhalte- und Überholweg haben einen festen Platz in der Fahrausbildung und Fahrerlaubnisprüfung. Aus diesem Grund sollte überprüft werden, ob dieser Stellenwert gerechtfertigt ist. Daher war zu testen, ob die Kenntnis dieser Regeln ein valides Kriterium für sicheres Fahren ist. Die Einschätzungsaufgaben zu den

Faustregeln haben sich in Vorstudie 1 nicht als valide und reliabel erwiesen. Daher gingen sie nicht in die Längsschnittstudie ein. Innerhalb der theoretischen Fahrerlaubnisprüfung wird die Kenntnis der Faustregeln teilweise mithilfe des Ergänzungsaufgabenformates explizit abgefragt. Um Wissen bezüglich der Faustregeln zu Anhalteweg, Sicherheitsabstand und Überholweg in der Längsschnittstudie zu erfassen, wurden zehn Regelaufgaben entwickelt. Diese unterschieden sich von den herkömmlichen Aufgaben dadurch, dass man, um die Aufgaben zu lösen, die Faustformeln nicht explizit formulieren musste. Die Aufgaben konnten auch durch erfahrungsbedingt richtiges Schätzen gelöst werden. Um dies zu ermöglichen, wurden die ursprünglichen Formelaufgaben der Fahrerlaubnisprüfung in ein Multiple-Choice-Aufgabenformat umgewandelt (vgl. Abbildung 5.1).

Sie fahren auf einer schmalen Straße. 20 m vor Ihnen betritt ein Fußgänger die Fahrbahn. Wann ist ein Zusammenprall unvermeidbar?

- 1) bei 40 km/h
- 2) bei 30 km/h
- 3) bei 20 km/h

Abbildung 5.1
Regelaufgabe Nr. 7

Experten sollte durch dieses Antwortformat ermöglicht werden, die richtige Lösung anhand ihrer Erfahrung zu schätzen, ohne die genaue Formel erinnern zu müssen. Novizen ohne Fahrerfahrung hingegen, sollten die Aufgaben nur lösen können, wenn sie die Faustformel bereits gelernt haben und diese richtig einsetzen.

5.2.2.2 Paper Folding Test

Alle Teilnehmer bearbeiteten eine Computerversion des *Paper Folding Tests* aus dem *Manual for Kit of Factor-Referenced Cognitive Tests* (Ekstrom et al., 1978). Eine Beschreibung des Tests wurde bereits in Kapitel 4.1.2 gegeben, worauf an dieser Stelle verwiesen werden soll.

5.2.2.3 Hidden Patterns Test

Zur zuverlässigen Erfassung räumlicher Fähigkeiten wurde der *Hidden Patterns Test*, ein weiteres Verfahren aus dem *Manual for Kit of Factor-Referenced Cognitive Tests* (Ekstrom et al., 1978) in der Längsschnittstudie eingesetzt. Er gehört zur Untergruppe der Tests zur flexiblen Wahrnehmung von Geschlossenheit (*Flexibility of Closure*). Der in Vorstudie 1 eingesetzte Paper Folding Test zur Erfassung der Visualisierungsfähigkeit sollte damit durch

einen Test ergänzt werden, der die räumlichen Fähigkeiten einer Person unter Einbezug einer Zeitkomponente misst. Es wurde angenommen, dass dies eher den Anforderungen der Reaktionszeitaufgaben, die in der Längsschnittstudie zum Einsatz kamen, sowie des realen Fahrens selbst entspricht.

Die Aufgabe der Teilnehmer besteht beim *Hidden Patterns Test* darin, eine graphische Strichfigur innerhalb eines Musters zu suchen und zu entscheiden, ob die Figur im jeweiligen Muster enthalten ist oder nicht. Jedes einzelne Item besteht aus einem geometrischen Muster, bei dem durch anklicken des „Ja“- oder „Nein“-Feldes angegeben werden soll, ob die Referenzfigur darin integriert ist. Die Figur, die gesucht werden soll, bleibt im gesamten Test gleich und wird am oberen rechten Rand des Tests abgebildet.

Der Test besteht aus zwei Teilen mit je 200 Mustern, die in Reihen auf zwei Seiten verteilt sind. Zur Bearbeitung besteht ein Zeitlimit von drei Minuten pro Seite. Im Instruktionstext, der vor der Bearbeitung der Items präsentiert wird, werden die Teilnehmer angewiesen gleichzeitig schnell und akkurat zu arbeiten. Die Gesamtpunktzahl des Tests ergibt sich aus der Anzahl richtig gelöster Items abzüglich der falsch gelösten Aufgaben. Die einzelnen Items können relativ schnell gelöst werden; In der Regel reicht die Zeit jedoch nicht aus, um alle Items zu bearbeiten. Für die Längsschnittstudie wurde eine Computerversion dieses Testverfahrens erstellt.

5.2.2.4 Persönliche Angaben

Die demographischen Daten der Teilnehmer wurden am ersten Messzeitpunkt abgefragt. Alle Probanden bearbeiteten am Computer Fragen zu Geschlecht, Alter, höchstem Bildungsabschluss, Berufstätigkeit und der Anzahl der Einwohner des eigenen Wohnortes. Jedem Teilnehmer wurden an jedem Messzeitpunkt Fragen bezüglich seiner Fahrschulausbildung, seines Fahrverhaltens und seiner bisherigen Unfallgeschichte zur Beantwortung am Computer vorgegeben. Welche Fragen dem Probanden jeweils gestellt wurden, war abhängig von Messzeitpunkt, Expertise und dem Stand der Fahrausbildung. Lediglich die Novizen wurden gebeten, Angaben zur Anzahl der bisher absolvierten Theorie- und Praxisstunden (auch Sonderfahrten) zu machen. Zudem sollten sie angeben, ob sie bereits die theoretische oder praktische Fahrerlaubnisprüfung abgelegt hatten und wie viele Versuche sie jeweils bis zum Bestehen benötigt hatten. Die Probanden wurden bzgl. ihrer Teilnahme am Begleiteten Fahren ab 17 Jahren und an Sicherheitstrainings befragt. Alle Teilnehmer sollten zudem angeben, ob sie eine Fahrerlaubnis für weitere Klassen besitzen.

Die Experten wurden an jedem Messzeitpunkt zu Unfällen unterschiedlicher Schweregrade, die sie bisher und in den letzten sechs Monaten verursacht hatten, befragt. Außerdem wurden sie darum gebeten, zu berichten, wie viele kritische Situationen im Verkehr sie in den letzten vier Wochen erlebt hatten. Sobald die Novizen angaben, ihre Fahrerlaubnisprüfung erfolgreich abgelegt zu haben, wurden auch ihnen die Fragen zu Unfällen und kritischen Situationen im Verkehr zur Bearbeitung dargeboten. Analog zum Vorgehen in Vorstudie 1 wurden zur Erfassung der kritischen Situationen und Unfälle Fragen eines von einer Projektgruppe um Debus, Leutner und Brünken (2008) entwickelten Fragebogens eingesetzt.

5.2.2.5 Sensation Seeking

In Vorstudie 2 wurden den Teilnehmern zur Erfassung der Sensationssuche 12 Items der Skala *Thrill and Adventure Seeking* in Anlehnung an Zuckermann (1979) vorgelegt, die wie bei Gniech, Oetting und Brohl (1993) auf einer fünfstufigen Ratingskala mit den Endpunkten +2 („trifft zu“) und -2 („trifft nicht zu“) zu beantworten waren. Es zeigten sich nur geringe Zusammenhänge zwischen den Angaben der Teilnehmer in dieser Skala und ihren Leistungen in den Reaktionszeitaufgaben. Es wurde angenommen, dass deutlichere Zusammenhänge möglicherweise nicht festgestellt wurden, weil sich die Aussagen in den Items nicht direkt auf das Autofahren beziehen. Daher wurden für die Längsschnittstudie zur Erfassung von Aspekten der Sensationssuche fünf Items zur verkehrsbezogenen Aggressivität aus dem „Verkehrsspezifischen Itempool“ (Schmidt, L., 1986) und neun Items zu Impulsivität und Sensationssuche aus dem Untersuchungsmaterial der australischen Drive-Studie (Ivers et al., 2006) eingesetzt. Das Antwortformat war dichotom: Bei Zustimmung zu einer Aussage sollte „ja“ angekreuzt werden, bei Ablehnung sollte die Antwort „nein“ gewählt werden. Die ausgewählten Items wurden in dieser Kombination auch in einer Studie von Debus et al. (2008) eingesetzt.

Die in der Längsschnittstudie eingesetzten Fragebögen sind in Anhang B verfügbar.

5.2.3 Durchführung

Bei der durchgeführten Studie handelte es sich um eine Längsschnittstudie mit drei Messzeitpunkten, die im Abstand von jeweils etwa vier Monaten stattfanden. Dazu wurden die Versuchspersonen entweder in die Räume der kooperierenden Fahr- und Berufsschulen oder in das Computerlabor der Fachrichtung Bildungswissenschaften der Universität des Saarlandes eingeladen. Die Ansprache von Experten und Novizen erfolgte über Aushänge und

Direktansprache in Fahrschulen im Saarland und Rheinland-Pfalz sowie an verschiedenen Berufsschulen im Saarland und an der Universität des Saarlandes. Zu jedem Messzeitpunkt wurde das Versuchsmaterial von den Teilnehmern am Computer bearbeitet. In der Regel wurden Gruppen von vier bis zehn Personen gleichzeitig getestet. Am dritten Messzeitpunkt wurde jeder Versuchsperson eine Aufwandentschädigung von 30 Euro ausgezahlt.

Zur Messung der Leistungsveränderung in den Wissens- und Reaktionszeitaufgaben wurden diese Aufgaben an jedem Messzeitpunkt zur Bearbeitung vorgegeben. Dabei wurde den Teilnehmern eine von drei Versionen (A, B, C) dieser Testaufgaben vorgelegt. Die Reihenfolge der Versionen wurde für jede Versuchsperson im Vorhinein zufällig festgelegt. Jeweils bevor die Wissens-, Regel und Reaktionszeitaufgaben bearbeitet werden mussten, wurden demographischen Angaben und Angaben zu Unfällen und kritischen Situationen erfragt. Am ersten Messzeitpunkt wurden zusätzlich die beiden Tests zu räumlichen Fähigkeiten am Computer bearbeitet, bevor die fahr- und verkehrsspezifischen Leistungstests vorgegeben wurden.

5.3 Ergebnisse

5.3.1 Fahrerfahrung und Fahrausbildung

Zu allen Messzeitpunkten wurden die Teilnehmer der Längsschnittstudie auch um einige Angaben bezüglich ihrer Fahrausbildung und ihrer bisherigen Erfahrung als Fahrer (nur die Experten) gebeten. Fast die Hälfte der Novizen nahm am Begleiteten Fahren ab 17 Jahren teil (46 %), während nur acht Experten daran teilgenommen hatten (13 %). Laut dem saarländischen Ministerium für Umwelt, Energie und Verkehr ist die Anzahl der Personen die am Begleiteten Fahren ab 17 teilgenommen haben, seit der Einführung dieser Regelung stetig gestiegen. Im Jahre 2009 waren es fast 39 % der Fahrschüler. Die Stichprobe dieser Studie erscheint, angesichts des steigenden Trends der Teilnehmer am BF 17, repräsentativ für die Grundgesamtheit saarländischer Fahrschüler.

Die meisten Novizen gaben zu Messzeitpunkt 1 an, zurzeit die Fahrschule zu besuchen (47 %) oder zumindest in den nächsten Monaten mit dem Fahrschulunterricht zu beginnen (33 %). Ein Großteil der Fahrnovizen hatte bis zum ersten Messzeitpunkt maximal drei praktische Fahrstunden (vgl. Tabelle 5.6). Die Anzahl der praktischen Fahrstunden stieg über die drei Messzeitpunkte an. Im Falle der Experten wurde erhoben, wie viele praktische Fahrstunden sie bis zum Erwerb ihrer Fahrerlaubnisprüfung absolviert hatten. Es zeigte sich, dass über die Hälfte der Experten (55 %) 16 oder mehr praktische Fahrstunden hatten.

Fast alle Novizen gaben an, bis zum ersten Messzeitpunkt noch keine Sonderfahrten, wie Überlandfahrten (90 %), Autobahnfahrten (98 %), Beleuchtungsfahrten (99 %) und Schlechtwetterfahrten (89 %), absolviert zu haben. Die Anzahl derjenigen Experten, die während ihrer Fahrausbildung nicht an Sonderfahrten wie Überlandfahrten (5 %), Autobahnfahrten (5 %) und Beleuchtungsfahrten (7 %) teilgenommen hatte, war demgegenüber sehr gering. Etwas mehr Experten gaben an, während ihrer Fahrausbildung nie eine Fahrstunde bei schlechtem Wetter absolviert zu haben (19 %).

Tabelle 5.6

Absolvierte Praktische Fahrstunden (%)

Anzahl praktischer Fahrstunden	Novizen			Experten
	MZP1 (n = 83)	MZP2 (n = 63)	MZP3 (n = 65)	
max. 3	76 (92)	31 (49)	23 (35)	3 (5)
4 bis 6	2 (2)	3 (5)	2 (3)	1 (2)
7 bis 9	4 (5)	6 (9.5)	2 (3)	3 (5)
10 bis 12	1 (1)	5 (8)	4 (6)	7 (11)
13 bis 15	-	3 (5)	1 (2)	6 (10)
16 bis 18	-	2 (3)	-	4 (6)
19 bis 21	-	1 (2)	-	10 (16)
22 bis 24	-	6 (9.5)	15 (23)	14 (23)
25 bis 27	-	2 (3)	18 (28)	4 (6)
> 30	-	4 (6)	-	2 (3)
keine Angabe	-	-	-	8 (13)

Die Anzahl der bisher absolvierten Theoriestunden wurde lediglich am ersten Messzeitpunkt von den Novizen erfragt. Es zeigte sich, dass die Mehrzahl der Personen (43 %) bis zum ersten Messzeitpunkt erst maximal dreimal am Theorieunterricht teilgenommen hatte. Etwa ein Viertel der Novizen hatten bereits die Pflichtstundenanzahl von 14 Doppelstunden abgeleistet (24 %).

Die meisten Experten (69 %) gaben an, neben dem PKW-Führerschein keine weitere Fahrerlaubnis zu besitzen. Einige Personen besaßen jedoch eine Fahrerlaubnis für Motorrad, LKW oder land- und forstwirtschaftliche Zugmaschinen.

Etwa einem Dreiviertel der Experten (76 %) stand ein eigener oder geliehener PKW zur Verfügung. Die Fahrerfahrung der Experten wurde über zwei Angaben erfasst. Zum einen sollten die Experten angeben, seit wann sie eine Fahrerlaubnis der Klasse B besaßen und zum anderen, wie viele Kilometer sie seitdem selbstständig gefahren waren. Es zeigte sich, dass die Personen zwischen drei und 35 Jahren im Besitz einer Fahrerlaubnis waren. Im Mittel besaßen die Experten seit 8.41 ($SD = 7.22$) Jahren einen Führerschein. Seit dem Erwerb ihrer Fahrerlaubnis waren zwei Drittel der Experten (66 %) bisher über 10 000 Kilometer selbstständig gefahren. Aus Tabelle 5.7 wird ersichtlich, dass die Fahrleistung der Experten sich über die drei Messzeitpunkte nicht deutlich veränderte. Die Fahrleistung der Novizen kann erst ab dem zweiten Messzeitpunkt angegeben werden, da erst ab diesem Zeitpunkt einige Novizen ihre Fahrerlaubnis erworben hatten und bereits selbstständig gefahren waren.

Tabelle 5.7

Fahrerfahrung der Teilnehmer in km (%)

Fahrerfahrung in km	Novizen ^{a)}			Experten	
	MZP2 (n = 12)	MZP3 (n = 30)	MZP1 (n = 62)	MZP2 (n = 57)	MZP3 (n = 55)
max. 500	9 (75)	8 (27)	2 (3)	1 (2)	2 (4)
501-2000	3 (25)	12 (40)	3 (5)	3 (5)	2 (4)
2001-5000	-	7 (23)	7 (11)	3 (5)	5 (9)
5001-10 000	-	2 (7)	9 (15)	7 (12)	2 (4)
10 001-50 000	-	1 (3)	12 (19)	17 (30)	15 (27)
50 001-100 000	-	-	7 (11)	10 (18)	12 (22)
100 001-250 000	-	-	14 (23)	9 (16)	5 (9)
250 001-500 000	-	-	6 (10)	6 (10)	9 (16)
500 001-1 000 000	-	-	2 (3)	1 (2)	3 (5)

Anmerkung. ^{a)}Es wird nur die Fahrerfahrung derjenigen Novizen aufgeführt, die bis zum jeweiligen Messzeitpunkt eine Fahrerlaubnis erworben haben.

Die Experten wurden an jedem Messzeitpunkt gefragt, wie viele Kilometer sie durchschnittlich in einer Woche als Fahrer zurücklegen. Auch den Novizen wurde diese Frage gestellt, sobald diese die Fahrerlaubnis erworben hatten. Die Experten gaben an allen drei Messzeitpunkten insgesamt an, wöchentlich maximal 750 Kilometer zu fahren (vgl. Tabelle 5.8).

Tabelle 5.8

Durchschnittliche wöchentliche Fahrleistung der Teilnehmer (%)

Wöchentliche Fahrleistung in km	Novizen ^{a)}			Experten	
	MZP2 (n = 12)	MZP3 (n = 30)	MZP1 (n = 62)	MZP2 (n = 57)	MZP3 (n = 55)
max. 50	7 (59)	13 (43)	17 (27)	18 (31)	16 (29)
51-100	3 (25)	6 (20)	7 (11)	6 (10.5)	7 (13)
101-150	1 (8)	2 (7)	5 (8)	6 (10.5)	9 (17)
151-200	1 (8)	3 (10)	13 (21)	8 (14)	3 (5)
201-250	-	2 (7)	6 (10)	6 (10.5)	3 (5)
251-300	-	1 (3)	3 (5)	3 (5)	6 (11)
301-350	-	3 (10)	5 (8)	6 (10.5)	4 (7)
351-400	-	-	5 (8)	1 (2)	3 (5)
401-450	-	-	-	1 (2)	1 (2)
451-500	-	-	1 (2)	1 (2)	1 (2)
501-550	-	-	-	-	-
551-600	-	-	-	1 (2)	1 (2)
601-650	-	-	-	-	-
651-700	-	-	-	-	-
701-750	-	-	-	-	1 (2)

Anmerkung. ^{a)}Es wird nur die Fahrerfahrung derjenigen Novizen aufgeführt, die bis zum jeweiligen Messzeitpunkt eine Fahrerlaubnis erworben haben.

Während von den Experten wenige Personen angaben, pro Woche über 400 Kilometer im Durchschnitt selbst zu fahren, berichtete keiner der Novizen über 350 Kilometer in der Woche als Fahrer zurückzulegen. Am häufigsten wurde von den Experten und Novizen die Kategorie bis maximal 50 Kilometer pro Woche gewählt.

Zu besonderen Gelegenheiten (z.B. Urlaub) gaben die Experten an, bis über 5 000 Kilometer selbst gefahren zu sein. Bei der relativ gleichmäßigen Verteilung der Häufigkeiten auf die verschiedenen vorgegebenen Kategorien gab es zwei kleine Spitzen an den Extremen. Ungefähr 13 % hatten bisher erst maximal 100 Kilometer zu besonderen Gelegenheiten zurückgelegt. Weitere 13 % waren zu besonderen Gelegenheiten bisher bereits über 5 000 Kilometer gefahren. Die Novizen, die eine Fahrerlaubnis erworben hatten, gaben an, bis zum letzten Messzeitpunkt maximal 2 000 Kilometer zu besonderen Gelegenheiten selbst gefahren zu sein. Am häufigsten wurde die Kategorie bis maximal 100 Kilometer gewählt.

Insgesamt gaben zehn Experten und ein Novize an, bereits ein Fahrsicherheitstraining absolviert zu haben.

5.3.2 Unfälle und kritische Situationen

Die Experten unter den Versuchsteilnehmern wurden zu jedem Messzeitpunkt gebeten, Angaben zu Unfällen, die sie in den letzten sechs Monaten und insgesamt seit dem Erwerb ihrer Fahrerlaubnis verursacht hatten, zu Verwarnungsgeldern, die sie in den letzten sechs Monaten und insgesamt bisher bezahlen mussten und zu kritischen Verkehrssituationen, die sie in den letzten vier Wochen erlebt hatten, zu machen. Auch die Novizen wurden um diese Daten gebeten, sobald sie ihre Fahrerlaubnis erworben hatten. In Tabelle 5.9 werden die von den Teilnehmern angegebenen durchschnittlichen Unfälle, Verwarnungsgelder und kritischen Verkehrssituationen aufgeführt.

Tabelle 5.9

Durchschnittliche Anzahl von selbstverschuldeten Unfällen, Verwarnungsgeldern und kritischen Situationen (Standardabweichung)

	Novizen		Experten		
	MZP2 (n = 12)	MZP3 (n = 30)	MZP1 (n = 62)	MZP2 (n = 57)	MZP3 (n = 55)
Unfälle gesamt	0.08 (0.29)	0.07 (0.25)	1.16 (1.32)	1.05 (1.33)	1.20 (1.54)
Unfälle (6 Monate)	0.08 (0.29)	0.07 (0.25)	0.08 (0.38)	0.11 (0.36)	0.04 (0.27)
Verwarnungsgelder gesamt		0.07 (0.25)	2.36 (2.59)	2.17 (2.26)	3.22 (3.60)
Verwarnungsgelder (6 Monate)		0.07 (0.25)	0.45 (0.78)	0.44 (0.78)	0.45 (0.94)
Kritische Situationen	1.83 (3.04)	3.3 (13.89)	1.52 (2.16)	1.44 (2.31)	0.98 (1.74)

Die Novizen berichteten am zweiten und dritten Messzeitpunkt, bisher ausschließlich Unfälle mit geringem oder mittelhohem Sachschaden verursacht zu haben. Einige Experten gaben hingegen auch an, bereits größere Unfälle mit Personenschaden verursacht zu haben. Bei den meisten Unfällen handelte es sich dabei um kleinere Unfälle, bei denen es nicht notwendig war, die Polizei zu verständigen.

Am häufigsten wurden kritische Situationen bei Nässe erlebt. 21 % der Studienteilnehmer gaben an, in den letzten vier Wochen eine kritische Verkehrssituation

erlebt zu haben, die sich aufgrund nasser Fahrbahn und Regen ergeben hatte. Sehr selten kamen Situationen vor, in denen beinahe Fußgänger zu Schaden gekommen wären. Diese Situation hatte nur eine einzige Person in den letzten vier Wochen erlebt.

5.3.3 Parallelität der Testformen

Zur Messung der Leistung in den unterschiedlichen verkehrsspezifischen Aufgabenformaten zu den drei Messzeitpunkten wurden zu jeder Aufgabe zusätzlich zwei Parallelaufgaben erstellt. Die Aufgaben wurden zu drei Paralleltestformen zusammengestellt, sodass von jeder Aufgabe eine Parallelversion in jeder Paralleltestform enthalten war.

Um festzustellen, ob sich die unterschiedlichen Testformen in ihrer Schwierigkeit unterschieden, wurden zu jedem Messzeitpunkt die Mittelwerte, die in den drei Paralleltestversionen in den unterschiedlichen Aufgabenformaten erreicht wurden, miteinander verglichen (vgl. Tabelle 5.10).

Tabelle 5.10

Parallelität der Testformen A, B und C: Mittelwerte (Standardabweichung) der Anzahl der richtigen Lösungen in den Wissens- und Reaktionszeitaufgaben

		MZP1	MZP2	MZP3
Wissensaufgaben	A	12.98 (4.99)	16.34 (4.06)	15.79 (3.44)
	B	14.14 (4.91)	15.95 (3.66)	15.11 (5.54)
	C	13.83 (5.10)	15.09 (4.80)	16.78 (2.82)
Reaktionszeitaufgaben	A	14.73 (4.82)	18.54 (4.90)	16.64 (6.00)
	B	15.13 (4.96)	16.60 (5.39)	15.61 (5.34)
	C	15.24 (5.69)	17.18 (4.50)	17.73 (4.85)

Es zeigten sich zu MZP1 keine statistisch bedeutsamen Unterschiede in Abhängigkeit von der Paralleltestversion für die Wissensaufgaben ($F(2, 132) = .64$; $p = .53$) und für die Lösungen bei den Reaktionszeitaufgaben ($F(2, 132) = .12$; $p = .89$). Auch zum zweiten Messzeitpunkt waren keine statistisch bedeutsamen Leistungsunterschiede zwischen den drei Paralleltestversionen bei den Wissensaufgaben ($F(2, 110) = .86$; $p = .45$) und für die Lösungen bei den Reaktionszeitaufgaben ($F(2, 110) = .159$; $p = .21$) feststellbar. Signifikante Leistungsunterschiede zwischen den drei Paralleltestversionen ließen sich auch zum dritten Messzeitpunkt für die Wissensaufgaben ($F(2, 108) = 1.63$; $p = .20$) und für die Lösungen bei den Reaktionszeitaufgaben ($F(2, 109) = 1.51$; $p = .23$) nicht feststellen. Die

Paralleltestformen waren demnach bezüglich ihrer Schwierigkeit zu allen drei Messzeitpunkten vergleichbar.

5.3.4 Kovariaten

Verarbeitung visueller Darstellungen

Zur Messung der Verarbeitung visueller Darstellungen der Teilnehmer wurden der Hidden Patterns Test (Flexibilität der Wahrnehmung von Geschlossenheit) und der Paper Folding Test (Visualisierungsfähigkeit) eingesetzt. Es zeigte sich, dass die Experten ($M = 91.06$; $SD = 30.80$) im Mittel mehr Aufgaben im Hidden Patterns Test richtig lösten als die Novizen ($M = 63.14$; $SD = 23.40$). Der Unterschied erwies sich als statistisch bedeutsam ($T(141) = 5.86$; $p \leq .001$). Auch im Paper Folding Test erreichten die Experten durchschnittlich höhere Leistungswerte ($M = 8.73$; $SD = 5.13$) als die Novizen ($M = 5.95$; $SD = 4.89$). Der Mittelwertunterschied hielt ebenfalls der Überprüfung auf statistische Signifikanz stand ($T(141) = 3.29$; $p \leq .001$).

Für die Gruppe der Experten zeigte sich zwischen den Leistungen im Hidden Patterns Test und im Paper Folding Test eine signifikante positive Korrelation von $r = .56$ ($p \leq .001$). Die Leistungen der Novizen in den beiden Tests zu räumlichen Fähigkeiten korrelierten ebenfalls signifikant zu $r = .36$ ($p = .004$) miteinander.

Aufgrund der festgestellten Leistungsunterschiede zwischen den Experten und den Novizen in den beiden Testverfahren zum Umgang mit grafischen Darstellungen, sollen die Visualisierungsfähigkeit sowie die Flexibilität der Wahrnehmung von Geschlossenheit (FWG) als Kovariaten in die Analysen eingehen. Um den Zusammenhang zwischen den beiden Kovariaten und den einzelnen verkehrsspezifischen Leistungsmaßen der Experten und Novizen zu spezifizieren, wurden bivariate Korrelationen zwischen den einzelnen Maßen berechnet, die in Tabelle 5.11 aufgeführt werden.

Insgesamt zeigten sich durchweg positive geringe bis mittelhohe Korrelationen zwischen den Leistungen in den Kovariaten und in den Wissens-, Regel- und Reaktionszeitaufgaben (Anzahl der richtigen Lösungen). Meist negative geringe bis mittelhohe Zusammenhänge zeigten sich zwischen den Leistungen in den beiden Kovariaten und den Reaktionszeiten. Insgesamt scheint demnach hohe Leistung in den Kovariaten mit hoher Leistung in den verkehrsspezifischen Leistungsmaßen einher zu gehen. Dabei fällt auf, dass die Korrelationen bei der Gruppe der Novizen insgesamt geringer ausfallen und seltener statistisch bedeutsam sind als bei den Experten.

Tabelle 5.11

Korrelative Zusammenhänge (Pearson r) zwischen den Leistungen im Paper Folding Test, Hidden Patterns Test und den Leistungen in den verkehrsspezifischen Leistungsmaßen

		Paper Folding Test			Hidden Patterns Test		
		Experten	Novizen	Gesamt	Experten	Novizen	Gesamt
MZP1	Wissensaufgaben	.31*	.22*	.28**	.18	.13	.20*
	Regelaufgaben	.34**	.15	.26**	.35**	-.01	.22**
	Reaktionstest (Anzahl)	.33*	.08	.25**	.47**	.23*	.44**
	Reaktionstest (Zeiten)	-.21	-.15	-.25**	-.27*	-.14	-.33**
MZP2	Wissensaufgaben	.12	.29*	.21*	.30*	.17	.17
	Regelaufgaben	.39**	.27*	.31**	.41**	.10	.23**
	Reaktionstest (Anzahl)	.30*	.06	.24**	.34*	.06	.30**
	Reaktionstest (Zeiten)	-.12	.09	-.11	-.22	-.22	-.33**
MZP3	Wissensaufgaben	.10	.23	.17	.06	.12	.06
	Regelaufgaben	.42**	.23	.29**	.16	.23	.16
	Reaktionstest (Anzahl)	.27*	.13	.29**	.18	-.01	.22
	Reaktionstest (Zeiten)	-.04	-.04	-.12	.02	.12	-.05

Anmerkungen. * $p < .05$; ** $p < .01$

Sensation Seeking

Am zweiten Messzeitpunkt wurde zur Überprüfung der Tendenz zur Sensationssuche eine Skala mit 11 Items vorgegeben. Es zeigte sich, dass die Experten im Mittel zwar einen geringeren Wert auf dieser Skala erreichten ($M = 4.47$; $SD = 2.92$) als die Novizen ($M = 4.67$; $SD = 2.40$), diese Differenz sich aber nicht als statistisch signifikant erwies ($T(83) = 0.31$; $p = .76$). Um mögliche Zusammenhänge zwischen den erreichten Werten auf der Sensation-Seeking-Skala und in den einzelnen verkehrsspezifischen Leistungsmaßen der Experten und Novizen zu den verschiedenen Messzeitpunkten zu identifizieren, wurden bivariate Korrelationen zwischen den einzelnen Maßen berechnet, die in Tabelle 5.12 aufgeführt werden. Es zeigten sich hauptsächlich geringe positive Korrelationen, die sich selten als statistisch signifikant erwiesen.

Aufgrund des fehlenden Unterschiedes zwischen den Experten und den Novizen in der Tendenz zur Sensationssuche, geht dieses Merkmal nicht als Kovariate in die weiteren Analysen ein.

Tabelle 5.12

Korrelative Zusammenhänge (Pearson r) zwischen den Ergebnissen in der Sensation-Seeking-Skala und den Leistungen in den verkehrsspezifischen Leistungsmaßen

		Sensation-Seeking-Skala		
		Experten	Novizen	Gesamt
MZP1	Wissensaufgaben	.26	.05	.18
	Regelaufgaben	.17	-.10	.07
	Reaktionstest (Anzahl)	.16	.01	.12
	Reaktionstest (Zeiten)	-.15	.39	-.03
MZP2	Wissensaufgaben	.30*	-.01	.19
	Regelaufgaben	.17	-.02	.11
	Reaktionstest (Anzahl)	.24	.01	.15
	Reaktionstest (Zeiten)	-.01	.04	.01
MZP3	Wissensaufgaben	.24	-.19	.13
	Regelaufgaben	.14	.30	.19
	Reaktionstest (Anzahl)	.34*	-.03	.22*
	Reaktionstest (Zeiten)	-.06	.01	-.06

Anmerkung. * $p < .05$

5.3.5 Ergebnisse der verkehrsspezifischen Leistungsmessung: Experten-Novizen-Vergleiche

Die innerhalb der Längsschnittstudie erhobenen Daten können zum einen querschnittlich betrachtet werden. Dabei werden in Experten-Novizen-Vergleichen die Leistungen von erfahrenen und unerfahrenen Fahrern zu jedem der drei Messzeitpunkte gegenübergestellt. Zum anderen können die Daten auch längsschnittlich untersucht werden, indem die Leistungsentwicklung der Teilnehmer, die das Testprogramm an allen drei Messzeitpunkten bearbeitet haben, über die Zeit betrachtet wird. Für die folgende Arbeit interessieren insbesondere die Ergebnisse des Experten-Novizen-Vergleichs zum ersten Messzeitpunkt, da dieser den Ausgangspunkt der Studie darstellt. Zu diesem Erhebungszeitpunkt sollten die expertisebedingten Leistungsunterschiede besonders deutlich ausfallen. Ein besonderes Augenmerk wird außerdem auf die Leistungsentwicklung derjenigen Personen gelegt, die an allen drei Erhebungen teilgenommen haben.

Im Folgenden werden zunächst die Ergebnisse der Experten-Novizen-Vergleiche zu jedem der drei Messzeitpunkte getrennt dargestellt. Zu jedem Messzeitpunkt werden die Ergebnisse bezüglich der Wissens-, Regel- und Reaktionszeitaufgaben aufgeführt. Dazu wurden jeweils Kovarianzanalysen mit den beiden Kovariaten Visualisierungsfähigkeit und FWG (Flexibilität der Wahrnehmung von Geschlossenheit) durchgeführt.

5.3.5.1 Messzeitpunkt 1

Die mittleren Summen der richtigen Lösungen, die die Novizen und Experten in den unterschiedlichen Aufgaben zu Messzeitpunkt 1 erreicht haben, werden in Tabelle 5.13 dargestellt. Auf deskriptivem Niveau zeigt sich, dass die Experten den Novizen zum ersten Messzeitpunkt in allen Aufgabentypen überlegen waren: Sie lösten mehr Wissens-, Regel- und Reaktionszeitaufgaben richtig und reagierten schneller auf potenzielle Gefahren in den Reaktionszeitaufgaben.

Tabelle 5.13

Mittelwerte und Standardabweichungen der Experten und Novizen in den Wissens-, Regel- und Reaktionszeitaufgaben zum ersten Messzeitpunkt

	Wissensaufgaben (Max. = 22)				Regelaufgaben (Max. = 10)				Reaktionstest			
					Anzahl korrekter Lösungen (Max. = 31)		Reaktionszeiten ^{a)}					
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>				
Novizen	12.70	(5.69)	4.73	(2.31)	13.50	(4.84)	.17	(0.51)				
Experten	14.29	(3.94)	5.63	(1.58)	16.69	(4.84)	-.19	(0.45)				

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung; *z* = mittlerer z-Wert; ^{a)}Reaktionszeiten wurden nur für die 18 *Hazard-Perception*- und die 7 Vorschriftaufgaben einbezogen.

Für die Leistung in den Wissensaufgaben als abhängige Variable zeigte sich in der Kovarianzanalyse kein Haupteffekt der Expertise ($F(1, 140) = 0.42$; $p = .52$). Es zeigte sich ein Haupteffekt der Kovariaten Visualisierungsfähigkeit ($F(1, 140) = 8.67$; $p = .004$; $\eta_p^2 = .06$) aber kein Haupteffekt der Kovariaten FWG ($F(1, 140) = 0.16$; $p = .69$).

Der Mittelwertunterschied zwischen den Novizen und den Experten bei den Regelaufgaben hielt der Überprüfung auf statistische Signifikanz in der Kovarianzanalyse ebenfalls nicht stand: Es konnte lediglich tendenziell ein Haupteffekt des Faktors Expertise festgestellt werden ($F(1, 140) = 3.38$; $p = .07$). Der Einfluss der Kovariate Visualisierungsfähigkeit war statistisch bedeutsam ($F(1, 140) = 5.35$; $p = .02$; $\eta_p^2 = .04$). Es lag kein Haupteffekt der Kovariate FWG vor ($F(1, 140) \leq .001$; $p \geq .99$).

Bei den Reaktionszeitaufgaben kann zwischen drei Aufgabentypen unterschieden werden: *Hazard-Perception*-Aufgaben, Vorschriftaufgaben und Distraktoren. In Vorstudie 2 hat sich gezeigt, dass die Art der Aufgaben mit dem Faktor Expertise interagiert. Daher sollte

auch bei der Längsschnittstudie der Messwiederholungsfaktor Aufgabentyp in die Analyse einbezogen werden. Da eine unterschiedliche Anzahl von Aufgaben den drei Aufgabentypen zuzuordnen war, wurden die korrekt gelösten Aufgaben an der jeweiligen Anzahl der Aufgaben des Aufgabentyps relativiert. In Tabelle 5.14 werden die statistischen Kennwerte getrennt nach den Aufgabentypen aufgeführt.

Tabelle 5.14

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen zum ersten Messzeitpunkt

	Distraktor		Hazard Perception		Vorschrift		gesamt	
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>
Novizen	.52	(.28)	.44	(.21)	.46	(.21)	.47	(.23)
Experten	.69	(.24)	.49	(.19)	.63	(.25)	.60	(.23)
Gesamt	.59	(.27)	.46	(.21)	.53	(.24)	.53	(.24)

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung

Die Kovarianzanalyse mit Messwiederholung zeigte einen signifikanten Haupteffekt des Faktors Expertise ($F(1, 139) = 5.20$; $p = .02$; $\eta_p^2 = .04$). Die Experten lösten anteilig an der Gesamtzahl der Aufgaben signifikant mehr Aufgaben als die Novizen. Der Effekt der Kovariate Visualisierungsfähigkeit war nicht statistisch bedeutsam ($F(1, 139) = 1.66$; $p = .20$). Es konnte jedoch ein Haupteffekt der Kovariate FWG festgestellt werden ($F(1, 139) = 9.20$; $p = .003$; $\eta_p^2 = .06$).

Für den Messwiederholungsfaktor Aufgabentyp hat sich gezeigt, dass Sphärizität nicht angenommen werden konnte ($p < .05$). Daher werden im Folgenden die nach Greenhouse-Geisser korrigierten Freiheitsgrade und p -Werte berichtet. Die Kovarianzanalyse zeigte keinen signifikanten Haupteffekt des Faktors Aufgabentyp ($F(2, 232) = 2.02$; $p = .14$). Es konnten auch keine signifikanten Interaktionseffekte zwischen Aufgabentyp und Expertise ($F(2, 232) = 2.32$; $p = .11$), zwischen Aufgabentyp und Visualisierungsfähigkeit ($F(2, 232) = 1.02$; $p = .35$) und zwischen Aufgabentyp und FWG ($F(2, 232) = 0.38$; $p = .65$) gezeigt werden.

Für die Analyse wurden die Reaktionszeiten am gemeinsamen Mittelwert aller Aufgaben z-standardisiert. Dadurch wurde gewährleistet, dass alle Aufgaben gleichermaßen

in das Ergebnis eingingen. Es wurde eine Kovarianzanalyse mit Messwiederholung auf dem Faktor Aufgabentyp durchgeführt. Der Messwiederholungsfaktor lag bei der Analyse der Reaktionszeiten zweifach gestuft vor, da Reaktionszeiten nur für die *Hazard-Perception*- und Vorschriftaufgaben, nicht aber für die Distraktoren, erfasst wurden. In Tabelle 5.15 werden die statistischen Kennwerte für die Reaktionszeiten getrennt nach den Aufgabentypen aufgeführt.

Tabelle 5.15

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: Reaktionszeiten) getrennt nach Aufgabentypen zum ersten Messzeitpunkt

	Hazard Perception		Vorschrift		gesamt	
	<i>z</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>
Novizen	.19	(.56)	.17	(.69)	.18	(.63)
Experten	-.26	(.54)	-.10	(.58)	-.18	(.56)
gesamt	-.01	(.60)	.05	(.66)	.02	(.61)

Anmerkungen. *z* = mittlerer *z*-Wert; *SD* = Standardabweichung

Es zeigte sich ein signifikanter Haupteffekt des Faktors Expertise ($F(1, 132) = 7.18; p = .008; \eta_p^2 = .05$). Die Experten reagierten schneller auf die Hinweise innerhalb der Videos als die Novizen. Die Kovarianzanalyse zeigte keine Haupteffekte der Kovariaten Visualisierungsfähigkeit ($F(1, 132) = 1.44; p = .23$) und FWG ($F(1, 132) = 1.28; p = .26$). Tendenziell konnte ein Haupteffekt des Messwiederholungsfaktors Aufgabentyp festgestellt werden. In den *Hazard-Perception*-Aufgaben wurde im Mittel schneller reagiert als in den Vorschriftaufgaben. Dieser Haupteffekt wird allerdings durch den Interaktionseffekt der Faktoren Expertise und Aufgabentyp ($F(1, 132) = 5.31; p = .02; \eta_p^2 = .04$) relativiert. Die Interaktion wird in Abbildung 5.2 veranschaulicht.

Zur genaueren Spezifikation der Interaktion wurden Scheffé-Tests für alle Einzelvergleiche durchgeführt. Dabei zeigten sich keine statistisch bedeutsamen Unterschiede zwischen den verschiedenen Bedingungen (alle $p > .05$).

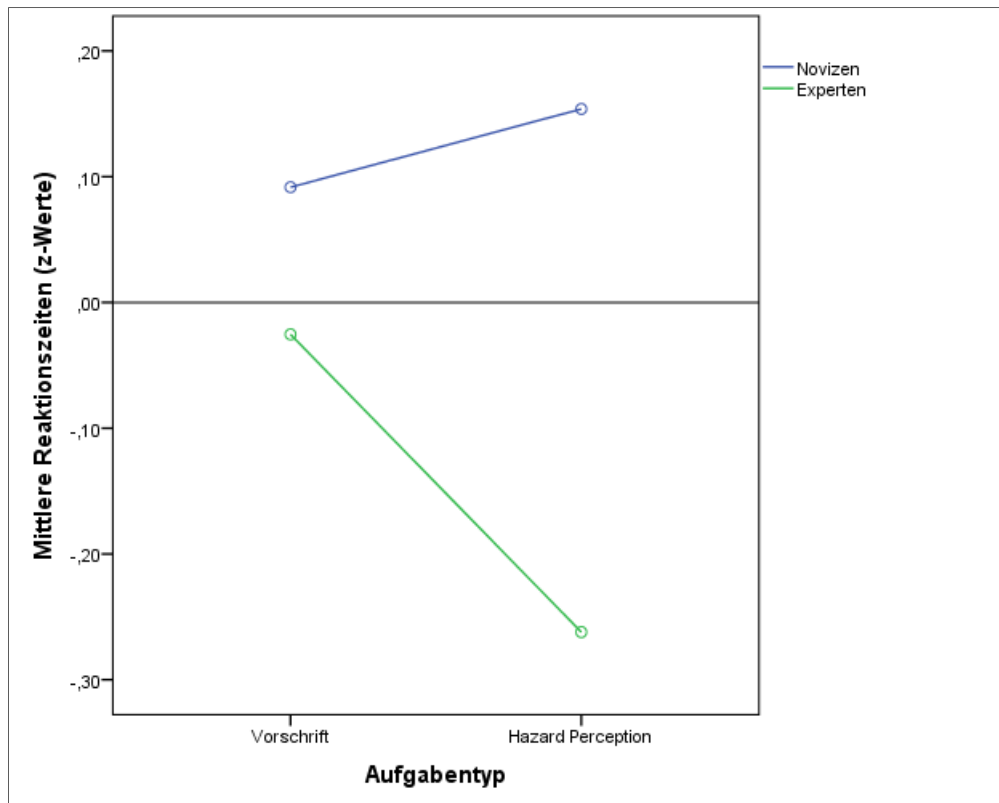


Abbildung 5.2

*Längsschnittstudie, MZP1: Ergebnisse zu Reaktionszeitaufgaben (z-Werte): Interaktion Aufgabentyp*Expertise*

Die Kovarianzanalyse zeigte keine signifikanten Interaktionen zwischen dem Faktor Aufgabentyp und den Kovariaten Visualisierungsfähigkeit ($F(1, 132) = 0.46; p = .50$) und FWG ($F(1, 132) = 1.02; p = .36$).

Insgesamt hat sich für den Messzeitpunkt 1 gezeigt, dass die Experten den Novizen in den Wissens- und Regelaufgaben nicht überlegen waren, diese aber bei den Reaktionszeitaufgaben deutlich übertrafen. Die erfahrenen Fahrer lösten mehr Reaktionszeitaufgaben richtig und reagierten schneller als die unerfahrenen Fahrer. Ein deutlicherer Unterschied in den Reaktionszeiten zwischen Experten und Novizen war zu verzeichnen, wenn es sich bei den Bremsgründen innerhalb der gezeigten Videos um explizite Gefahren handelte. Die Kovariate Visualisierungsfähigkeit erklärte nur bei den Wissens- und Regelaufgaben einen statistisch bedeutsamen Anteil an der Varianz, die Kovariate FWG hingegen nur im Reaktionstest bei der Anzahl korrekt identifizierter Bremshinweise. Die Zusammenhänge zwischen den Leistungen und den Kovariaten waren unabhängig von den experimentellen Bedingungen: Es zeigten sich keinerlei Wechselwirkungen zwischen den Kovariaten und den Faktoren Expertise bzw. Aufgabentyp.

5.3.5.2 Messzeitpunkt 2

In Tabelle 5.16 werden die mittleren Leistungen der Novizen und der Experten in den verschiedenen Aufgabenarten zum zweiten Messzeitpunkt dargestellt. Auf deskriptivem Niveau zeigt sich im Gegensatz zum ersten Messzeitpunkt keine generelle Überlegenheit der Experten gegenüber den Novizen mehr. Bei den Wissensaufgaben übertrafen die Novizen die Experten im Durchschnitt.

Tabelle 5.16

Mittelwerte und Standardabweichungen der Experten und Novizen in den Wissens-, Regel- und Reaktionszeitaufgaben zum zweiten Messzeitpunkt

	Wissensaufgaben (Max. = 22)		Regelaufgaben (Max. = 10)		Reaktionstest			
					Anzahl korrekter Lösungen (Max. = 31)		Reaktionszeiten ^{a)}	
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>
Novizen	15.22	(5.31)	5.40	(2.40)	15.42	(5.52)	0.09	(0.06)
Experten	15.17	(3.30)	5.78	(1.56)	18.56	(5.00)	- 0.10	(0.07)

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung; *z* = mittlerer z-Wert; ^{a)}Reaktionszeiten wurden nur für die 18 *Hazard-Perception*- und die 7 *Vorschrift*aufgaben einbezogen.

Auch für den zweiten Messzeitpunkt wurden getrennte Kovarianzanalysen für die verschiedenen Aufgabenarten berechnet. Am zweiten Messzeitpunkt nahmen auch Personen teil, die an der ersten Messung nicht teilgenommen hatten. Für diese Personen lagen daher keine Werte in den Kovariaten vor, die nur zum ersten Messzeitpunkt erhoben wurden. Um diese Personen dennoch in die Analyse einzubeziehen, wurden die fehlenden Werte durch den jeweiligen Mittelwert der Expertisegruppe, zu der die Person gehörte, imputiert.

Für die Wissensaufgaben zeigte sich weder ein signifikanter Haupteffekt der Expertise ($F(1, 122) = 2.05$; $p = .16$) noch der Kovariaten Visualisierungsfähigkeit ($F(1, 122) = 3.16$; $p = .08$) und FWG ($F(1, 122) = 1.84$; $p = .18$).

Für die Leistungen in den Regelaufgaben zeigte die Kovarianzanalyse keine signifikanten Haupteffekte der Expertise ($F(1, 122) = 0.45$; $p = .50$) und der Kovariate FWG ($F(1, 122) = 1.22$; $p = .27$). Es bestand allerdings ein statistisch bedeutsamer Haupteffekt der Kovariate Visualisierungsfähigkeit ($F(1, 122) = 7.39$; $p = .008$; $\eta_p^2 = .06$).

Die Ergebnisse für den Reaktionstest (abhängige Variable: korrekte Lösungen) werden in Tabelle 5.17 aufgeführt.

Tabelle 5.17

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen zum zweiten Messzeitpunkt

	Distraktor		Hazard Perception		Vorschrift		gesamt	
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>
Novizen	.62	(.29)	.51	(.24)	.48	(.23)	.54	(.25)
Experten	.74	(.24)	.58	(.22)	.58	(.22)	.63	(.23)
gesamt	.68	(.27)	.54	(.23)	.52	(.23)	.58	(.24)

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung

Die Kovarianzanalyse mit Messwiederholung auf dem dreifach gestuften Faktor Aufgabentyp zeigte keinen signifikanten Haupteffekt des Faktors Expertise ($F(1, 122) = 1.86$; $p = .18$) für die abhängige Variable korrekte Lösung der Reaktionszeitaufgaben. Es zeigten sich ebenfalls keine statistisch bedeutsamen Haupteffekte der Visualisierungsfähigkeit ($F(1, 122) = 0.91$; $p = .34$) und nur tendenziell der Kovariate FWG ($F(1, 122) = 3.58$; $p = .06$). Für den Messwiederholungsfaktor Aufgabentyp wurde die Sphäritätsannahme nicht bestätigt ($p < .05$). Daher werden im Folgenden die nach Greenhouse-Geisser korrigierten Freiheitsgrade und p-Werte aufgeführt. Die Kovarianzanalyse zeigte keinen signifikanten Haupteffekt des Faktors Aufgabentyp ($F(2, 220) = 0.85$; $p = .43$). Es konnten auch keine signifikanten Interaktionseffekte zwischen Aufgabentyp und Expertise ($F(2, 220) = 0.16$; $p = .83$), zwischen Aufgabentyp und Visualisierungsfähigkeit ($F(2, 220) = 0.54$; $p = .57$) und zwischen Aufgabentyp und FWG ($F(2, 220) = 0.12$; $p = .87$) gezeigt werden.

Die Ergebnisse für den Reaktionstest (abhängige Variable: Reaktionszeiten) werden in Tabelle 5.18 dargestellt.

Tabelle 5.18

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: Reaktionszeiten) getrennt nach Aufgabentypen zum zweiten Messzeitpunkt

	Hazard Perception		Vorschrift		gesamt	
	<i>z</i>	(<i>SD</i>)	<i>z</i>	(<i>SD</i>)	<i>z</i>	(<i>SD</i>)
Novizen	.13	(.50)	.14	(.67)	.14	(.59)
Experten	-.21	(.52)	-.12	(.71)	-.17	(.62)
gesamt	-.02	(.54)	.03	(.70)	-.01	(.61)

Anmerkungen. *z* = mittlerer *z*-Wert; *SD* = Standardabweichung

Für die Reaktionszeiten zeigte sich ein tendenzieller Haupteffekt des Faktors Expertise ($F(1, 113) = 3.19; p = .08$). Die Experten reagierten tendenziell schneller als die Novizen. Es zeigte sich kein Haupteffekt der Visualisierungsfähigkeit ($F(1, 113) = 2.52; p = .12$) aber ein statistisch bedeutsamer Effekt der Kovariate FWG ($F(1, 113) = 9.04; p = .003; \eta_p^2 = .07$). Die Kovarianzanalyse deckte keinen Haupteffekt des bei dieser Analyse zweifach gestuften Messwiederholungsfaktors Aufgabentyp auf ($F(1, 113) = 0.003; p = .96$). Es zeigten sich weder statistisch bedeutsame Wechselwirkungseffekte zwischen Aufgabentyp und Expertise ($F(1, 113) = 0.11; p = .74$), noch zwischen Aufgabentyp und Visualisierungsfähigkeit ($F(1, 113) = 0.008; p = .93$) und Aufgabentyp und FWG ($F(1, 113) = 0.04; p = .84$).

Für den Messzeitpunkt 2 hat sich insgesamt gezeigt, dass die Experten die Novizen weder in den Wissens-, noch in den Regelaufgaben oder bei der Anzahl der richtig identifizierten Bremshinweise übertrafen. Die Experten reagierten auch nur tendenziell schneller als die Novizen auf potenzielle Gefahren. Der Messwiederholungsfaktor Aufgabentyp hatte keinen Einfluss auf die Leistung im Reaktionstest. Die Kovariate Visualisierungsfähigkeit erklärte nur bei den Regelaufgaben einen bedeutsamen Anteil der Varianz. Die Kovariate FWG erklärte demgegenüber nur beim Reaktionstest einen signifikanten Varianzanteil. Es konnten keine Wechselwirkungen zwischen den Kovariaten und den Faktoren Expertise bzw. Aufgabentyp festgestellt werden. Zusammenhänge mit den Kovariaten waren demnach unabhängig von den experimentellen Bedingungen.

5.3.5.3 Messzeitpunkt 3

In Tabelle 5.19 werden die mittleren Leistungen, die die Novizen und die Experten in den unterschiedlichen Aufgabenformaten am dritten Messzeitpunkt erreicht haben, aufgeführt.

Tabelle 5.19

Mittelwerte und Standardabweichungen der Experten und Novizen in den Wissens-, Regel- und Reaktionszeitaufgaben zum dritten Messzeitpunkt

	Wissensaufgaben (Max. = 22)				Regelaufgaben (Max. =10)				Reaktionstest			
									Anzahl korrekter Lösungen (Max. = 31)		Reaktionszeiten ^{a)}	
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>				
Novizen	15.55	(5.38)	5.57	(2.21)	15.02	(5.50)	0.09	(0.54)				
Experten	15.67	(2.98)	5.73	(1.95)	18.07	(4.94)	-0.10	(0.48)				

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung; *z* = mittlerer z-Wert; ^{a)}Reaktionszeiten wurden nur für die 18 *Hazard-Perception*- und die 7 *Vorschriftaufgaben* einbezogen

Analog zum Vorgehen bei den ersten beiden Messzeitpunkten wurden auch für den dritten Messzeitpunkt getrennte Kovarianzanalysen für die verschiedenen Aufgabenarten berechnet.

Die Kovarianzanalyse zeigte keinen signifikanten Haupteffekt des Faktors Expertise ($F(1, 114) = 1.42$; $p = .24$) auf die Leistung in den Wissensaufgaben. Es zeigte sich tendenziell ein Haupteffekt der Kovariate Visualisierungsfähigkeit ($F(1, 114) = 3.61$; $p = .06$). Die Kovariate FWG hatte keinen Haupteffekt auf die Leistung in den Wissensaufgaben ($F(1, 114) = 0.02$; $p = .89$).

Für die Regelaufgaben deckte die Kovarianzanalyse einen tendenziellen Haupteffekt des Faktors Expertise ($F(1, 114) = 3.10$; $p = .08$) und einen statistisch bedeutsamen Haupteffekt der Kovariate Visualisierungsfähigkeit ($F(1, 114) = 8.70$; $p = .004$; $\eta_p^2 = .07$) auf. Es zeigte sich kein signifikanter Haupteffekt der Kovariate FWG auf die Leistung in den Regelaufgaben ($F(1, 114) = 0.59$; $p = .44$).

In Tabelle 5.20 werden die deskriptiven Kennwerte für den Reaktionstest (AV: korrekte Lösungen) getrennt nach den Aufgabentypen aufgeführt.

Tabelle 5.20

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen zum dritten Messzeitpunkt

	Distraktor		Hazard Perception		Vorschrift		gesamt	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Novizen	.64	(.29)	.48	(.23)	.47	(.26)	.53	(.26)
Experten	.75	(.24)	.57	(.20)	.58	(.26)	.63	(.23)
gesamt	.69	(.22)	.52	(.22)	.52	(.26)	.59	(.24)

Anmerkungen. M = Mittelwert; SD = Standardabweichung

Für die abhängige Variable korrekte Lösungen der Reaktionszeitaufgaben deckte die Kovarianzanalyse mit Messwiederholung einen tendenziellen Haupteffekt des Faktors Expertise auf ($F(1, 114) = 3.61$; $p = .06$) auf. Die Fahrexperten lösten anteilig an der Gesamtzahl der Aufgaben tendenziell mehr Aufgaben richtig als die Novizen. Es zeigte sich zudem ein tendenziell statistisch bedeutsamer Haupteffekt der Kovariate Visualisierungsfähigkeit ($F(1, 114) = 3.66$; $p = .06$). Ein signifikanter Haupteffekt der Kovariate FWG zeigte sich jedoch nicht ($F(1, 114) = 0.03$; $p = .85$). Für den Messwiederholungsfaktor Aufgabentyp wurde die Sphäritätsannahme nicht bestätigt ($p < .05$). Im Folgenden werden daher die nach Greenhouse-Geisser korrigierten Freiheitsgrade und p-Werte berichtet. Die Kovarianzanalyse zeigte einen statistisch bedeutsamen Haupteffekt des Aufgabentyps ($F(2, 186) = 9.30$; $p \leq .001$; $\eta_p^2 = .08$), der in Abbildung 5.3 veranschaulicht wird. Es wird deutlich, dass die Distraktoren häufiger richtig gelöst wurden als die *Hazard-Perception*- und die Vorschriftaufgaben. Paarweise Vergleiche (korrigiert nach Bonferroni) zeigten signifikante Differenzen zwischen den mittleren Lösungswahrscheinlichkeiten für die Distraktoren und die *Hazard-Perception*-Aufgaben (mittlere Differenz = .18; $p \leq .001$) sowie zwischen den Lösungswahrscheinlichkeiten für die Distraktoren und die Vorschriftaufgaben (mittlere Differenz = .18; $p \leq .001$). Der Unterschied zwischen den mittleren Lösungswahrscheinlichkeiten der *Hazard-Perception*- und den Vorschriftaufgaben erwies sich hingegen nicht als statistisch signifikant (mittlere Differenz = .002; $p = 1.00$).

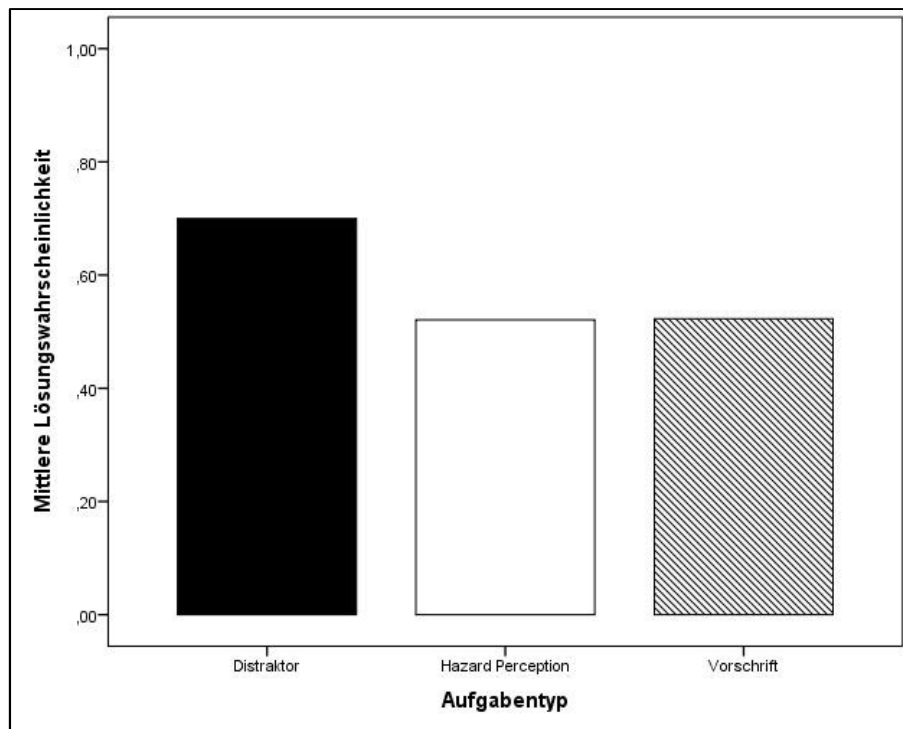


Abbildung 5.3

Ergebnisse zu Reaktionszeitaufgaben (Lösungswahrscheinlichkeit) zu Messzeitpunkt 3: Haupteffekt Aufgabentyp

Es konnten keine signifikanten Interaktionseffekte zwischen Aufgabentyp und Expertise ($F(2, 186) = 1.55$; $p = .22$) sowie zwischen Aufgabentyp und der Kovariate FWG ($F(2, 186) = 0.34$; $p = .67$) festgestellt werden. Zwischen Aufgabentyp und Visualisierungsfähigkeit zeigte sich allerdings eine signifikante Interaktion ($F(2, 186) = 5.60$; $p = .007$; $\eta_p^2 = .05$). Um diese Interaktion genauer zu klassifizieren wurden die Korrelationen zwischen Visualisierungsfähigkeit und Leistung im Reaktionstest getrennt für die drei Aufgabentypen (Distraktor, *Hazard Perception* und Vorschrift) berechnet. Die Korrelationskoeffizienten wurden für die wechselseitigen Vergleiche der Zusammenhänge zunächst in Fishers Z-Werte umgewandelt (vgl. Tabelle 5.21).

Tabelle 5.21

Zusammenhänge zwischen der Visualisierungsfähigkeit und der Leistung in den verschiedenen Aufgabentypen des Reaktionstests ($N = 120$)

	$r_{\text{(Pearson)}}$	Z_r	SE_{Z_r}
Distraktor	.03	.03	.09
Aufgabentyp Hazard Perception	.23*	.23	.09
Vorschrift	.37**	.39	.09

Anmerkungen. SE = Standardfehler; * $p < .05$; ** $p < .01$

Anschließend wurden gemäß Bortz (1999) als statistische Prüfgrößen z-Werte berechnet und auf statistische Signifikanz geprüft. Es zeigte sich, dass die Korrelation zwischen Visualisierungsfähigkeit und Leistung in den Distraktoraufgaben nicht signifikant geringer war als der korrelative Zusammenhang zwischen Visualisierungsfähigkeit und der Leistung in den *Hazard-Perception*-Aufgaben ($z = 1.56$; $p > .05$). Allerdings erwies sich der Unterschied zwischen der Korrelation von Visualisierungsfähigkeit und Leistung in den Distraktoraufgaben als signifikant geringer als die Korrelation von Visualisierungsfähigkeit und Leistung in den Vorschriftaufgaben ($z = 10.58$; $p \leq .001$). Die Korrelationen von Visualisierungsfähigkeit und Leistung unterschieden sich nicht signifikant zwischen den beiden Aufgabentypen *Hazard Perception* und Vorschrift ($z = 1.55$; $p > .05$).

Für die Reaktionszeiten, deren Kennwerte in Tabelle 5.22 getrennt nach Aufgabentypen aufgeführt sind, zeigte sich, dass die Experten tendenziell schneller reagierten als die Novizen ($F(1, 108) = 3.23$; $p = .08$). Es zeigte sich weder ein Haupteffekt der Visualisierungsfähigkeit ($F(1, 108) = 0.01$; $p = .95$) noch ein signifikanter Haupteffekt der Kovariate FWG ($F(1, 108) = 0.07$; $p = .79$). Die Kovarianzanalyse deckte keinen Haupteffekt des Messwiederholungsfaktors Aufgabentyp auf ($F(1, 108) = 0.05$; $p = .83$). Es zeigten sich auch weder statistisch bedeutsame Wechselwirkungseffekte zwischen Aufgabentyp und Expertise ($F(1, 108) = 0.02$; $p = .89$), noch zwischen Aufgabentyp und Visualisierungsfähigkeit ($F(1, 108) = 0.52$; $p = .47$) oder Aufgabentyp und FWG ($F(1, 108) = 0.01$; $p = .94$).

Tabelle 5.22

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: Reaktionszeiten) getrennt nach Aufgabentypen zum dritten Messzeitpunkt

	Hazard Perception		Vorschrift		gesamt	
	<i>z</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>	<i>z</i>	<i>(SD)</i>
Novizen	.05	(.58)	.12	(.66)	.09	(.62)
Experten	-.19	(.60)	-.04	(.68)	-.12	(.64)
gesamt	-.07	(.60)	.04	(.67)	-.02	(.63)

Anmerkungen. *z* = mittlerer z-Wert; *SD* = Standardabweichung

Zu Messzeitpunkt 3 übertrafen die Experten die Novizen in den Wissensaufgaben nicht und in den Regelaufgaben nur tendenziell. Im Reaktionstest beantworteten die Experten tendenziell mehr Aufgaben richtig und reagierten auch nur tendenziell schneller als die Novizen auf potenzielle Gefahren. Es wurden im Reaktionstest mehr Aufgaben richtig gelöst, wenn kein Bremshinweis auftauchte, als wenn es einen impliziten oder expliziten Bremsgrund gab. Der Faktor Aufgabentyp wirkte sich allerdings nicht bedeutsam auf die Reaktionszeiten aus. Die Kovariate Visualisierungsfähigkeit erklärte bei den Regelaufgaben und im Reaktionstest (Anzahl der korrekt gelösten Aufgaben) einen bedeutsamen Anteil der Varianz. Es konnten keine Wechselwirkungen zwischen den Kovariaten und den Faktoren Expertise bzw. Aufgabentyp für die Leistungen in den Wissens- und Regelaufgaben, sowie bei den Reaktionszeiten festgestellt werden. Zusammenhänge mit den Kovariaten waren demnach bei diesen Aufgabenformaten unabhängig von den experimentellen Bedingungen. Demgegenüber konnte jedoch festgestellt werden, dass die Höhe des Zusammenhangs zwischen der Leistung im Reaktionstest und der Visualisierungsfähigkeit abhängig vom Aufgabentyp variierte. Besonders deutliche positive Zusammenhänge zeigten sich zwischen der Visualisierungsfähigkeit und der Leistung in den Vorschriftaufgaben, während kein Zusammenhang zwischen der Kovariaten und der Leistung in den Distraktoraufgaben bestand.

5.3.6 Ergebnisse der verkehrsspezifischen Leistungsmessung:

Leistungsentwicklung über die Zeit

Aus den Ergebnissen der Experten-Novizen-Vergleiche wird bereits ersichtlich, dass die Experten zum einen nicht zu jedem Messzeitpunkt in sämtlichen Aufgabenarten den Novizen überlegen waren: In den Wissens- und Regelaufgaben übertrafen die Experten die Novizen kaum. Zum anderen fällt auf, dass die Effektstärken, die die Unterschiede zwischen den Experten und den Novizen in den Reaktionszeitaufgaben charakterisieren, über die Zeit geringer zu werden scheinen. Dies weist darauf hin, dass sich die Leistungen der Experten und der Novizen aneinander anpassen. Wie sich die Leistungsentwicklung in den unterschiedlichen Aufgabenformaten über die Zeit darstellt, soll im Folgenden überprüft werden. Dazu werden Varianzanalysen mit der abhängigen Variable Leistung in einem Aufgabenformat und den unabhängigen Variablen Zeit (messwiederholter Faktor, dreifach gestuft: MZP1 vs. MZP2 vs. MZP3) und Expertise (zweifach gestuft: Experten vs. Novizen) gerechnet. In die Analysen gehen nur die Daten derjenigen Personen ein, die an allen drei

Messzeitpunkten teilgenommen haben. Insgesamt haben 100 Personen (50 % Novizen) alle drei Messungen absolviert. Die Geschlechterverteilung war für beide Expertisegruppen gleich (59 % Frauen).

5.3.6.1 Kovariaten

Für die kleinere Stichprobe derjenigen Personen, die an allen drei Messzeitpunkten anwesend waren, wurde, analog zum Vorgehen für die Experten-Novizen-Vergleiche, zu den einzelnen Messzeitpunkten, der Einfluss möglicher Kovariaten überprüft.

Der T-Test für unabhängige Stichproben zeigte, dass sich die Experten und die Novizen in ihrer Visualisierungsfähigkeit signifikant unterschieden ($T(98) = 3.8$; $p \leq .001$; $d = .77$). Es zeigte sich, dass die Experten ($M = 9.27$; $SD = 4.70$) die Novizen ($M = 5.50$; $SD = 5.10$) im Paper Folding Test übertrafen. Auch in der Kovariate FWG unterschieden sich die beiden Expertisegruppen signifikant ($T(98) = 5.46$; $p \leq .001$; $d = 1.09$). Die Experten ($M = 93.38$; $SD = 31.76$) zeigten bessere Leistungen im Hidden Patterns Test als die Novizen ($M = 62.78$; $SD = 23.77$).

Für die Experten zeigte sich zwischen den Leistungen im Hidden Patterns Test und im Paper Folding Test eine signifikante Korrelation von $r = .54$ ($p \leq .001$). Die Leistungen der Novizen in den beiden Tests korrelierten ebenfalls signifikant zu $r = .31$ ($p = .02$) miteinander.

Aufgrund der festgestellten Leistungsunterschiede zwischen den beiden Expertisegruppen, gehen Visualisierungsfähigkeit sowie Flexibilität der Wahrnehmung von Geschlossenheit als Kovariaten in die Analysen ein.

Experten ($M = 5.66$; $SD = 2.61$) und Novizen ($M = 5.40$; $SD = 2.35$) unterschieden sich nicht signifikant hinsichtlich ihrer Werte auf der *Sensation-Seeking*-Skala ($T(98) = 0.42$; $p = .68$). *Sensation Seeking* wird daher nicht als Kovariate in die Analysen einbezogen.

5.3.6.2 Verkehrsspezifische Leistungsmaße

In Tabelle 5.23 werden die durchschnittlichen Leistungen der Experten und Novizen in den verschiedenen Aufgabenformaten deskriptiv dargestellt. Die Berechnung der z-standardisierten Werte für die Reaktionszeiten erfolgte auf unterschiedliche Arten. Daher werden diese im entsprechenden Abschnitt genauer dargestellt.

Tabelle 5.23

Deskriptive Ergebnisse der Längsschnittuntersuchung (N = 100) (vgl. Malone et al., 2012)

Zeitpunkt	Aufgabenart	Expertise			
		Novizen (n = 50)		Experten (n = 50)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
MZP1	Wissensaufgaben	13.54	5.92	14.94	3.22
	Regelaufgaben	5.06	2.49	5.68	1.60
	Reaktionstest (korrekte Lösungen)	14.56	4.59	17.62	4.37
MZP2	Wissensaufgaben	15.98	5.29	15.20	3.18
	Regelaufgaben	5.62	2.34	5.78	1.61
	Reaktionstest (korrekte Lösungen)	15.92	5.90	18.88	4.51
MZP3	Wissensaufgaben	16.60	4.90	15.76	2.73
	Regelaufgaben	5.92	1.89	5.62	1.97
	Reaktionstest (korrekte Lösungen)	15.32	5.73	18.2	4.75

Anmerkungen. *M* = Mittelwert ; *SD* = Standardabweichung

Es wurden für die unterschiedlichen Aufgabenarten Kovarianzanalysen mit Messwiederholung auf dem Faktor Zeit, der unabhängigen Variable Expertise und den beiden Kovariaten Visualisierungsfähigkeit und FWG durchgeführt.

In Abbildung 5.4 wird veranschaulicht, wie sich die Leistung der Experten und Novizen in den Wissensaufgaben über die drei Messzeitpunkte hinweg verändert. Die Ergebnisse der Kovarianzanalyse werden in Tabelle 5.24 zusammengefasst.

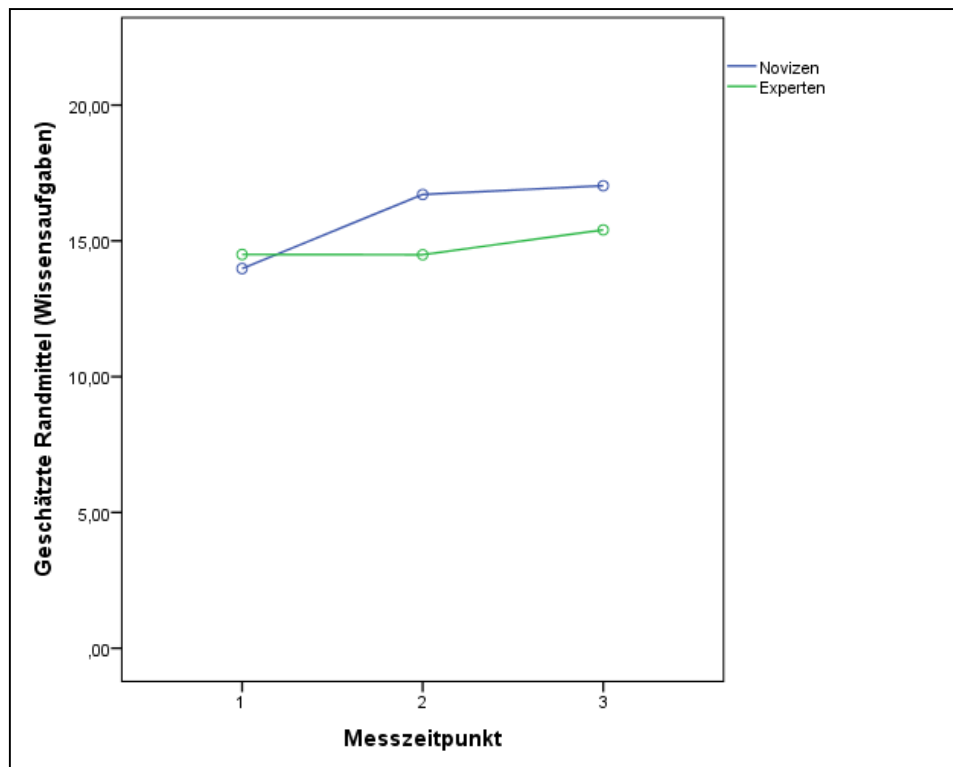


Abbildung 5.4

Leistungsentwicklung der Experten und Novizen in den Wissensaufgaben über die drei Messzeitpunkte hinweg (vgl. Malone et al., 2012)

Tabelle 5.24

Ergebnisse der Kovarianzanalyse für die Wissensaufgaben

Quelle der Varianz	<i>df</i>	<i>F</i>	<i>p</i>
Gruppenvergleiche			
Expertise (E)	1	1.54	.22
Visualisierungsfähigkeit (V)	1	2.48	.12
FWG (F)	1	1.29	.26
Fehler	94	(42.28)	
Messwiederholungsvergleiche ^{a)}			
Zeit (Z)	2	1.69	.19
Z x E	2	5.11	.01
Z x V	2	0.73	.46
Z x F	2	1.21	.29
Fehler	150	(8.97)	

Anmerkung. ^{a)}Für den Messwiederholungsfaktor Zeit konnte die Sphäritätsannahme nicht bestätigt werden ($p < .05$). Daher werden nach Greenhouse-Geisser korrigierte Freiheitsgrade und p-Werte berichtet.

Für die Wissensaufgaben zeigte sich kein signifikanter Haupteffekt der Expertise. Weder für die Visualisierungsfähigkeit noch für die Kovariate FWG konnten signifikante Haupteffekte festgestellt werden. Der Messwiederholungsfaktor Zeit hatte keinen statistisch bedeutsamen Haupteffekt auf die Leistung in den Wissensaufgaben. Allerdings lag ein signifikanter Interaktionseffekt zwischen den Faktoren Expertise und Zeit vor ($\eta_p^2 = .05$). Aus Abbildung 5.4 wird ersichtlich, dass die Differenz zwischen den Experten und Novizen sich über die Zeit verändert. Post-Hoc-Mehrfachvergleiche (Anpassung des Signifikanzniveaus nach Bonferroni) zeigten, dass die Leistung der Teilnehmer zu Messzeitpunkt 1 signifikant geringer war als zu Messzeitpunkt 2 (mittlere Differenz = 1.36; $p < .001$) und zu Messzeitpunkt 3 (mittlere Differenz = 1.98; $p < .001$). Zwischen Messzeitpunkt 2 und Messzeitpunkt 3 gab es keinen statistisch bedeutsamen Unterschied (mittlere Differenz = 0.62; $p = .19$). Im Scheffé-Test zeigte sich, dass diese Unterschiede lediglich auf die Entwicklung der Leistung der Novizen zurück zu führen war. Nur für die Novizen konnten signifikante Leistungssteigerungen zwischen dem ersten und dem zweiten Messzeitpunkt ($D = 2.73$; $\text{Diff}_{\text{krit}} = 2.69$; $p < .05$) und dem ersten und dem dritten Messzeitpunkt ($D = 3.05$; $\text{Diff}_{\text{krit}} = 2.69$; $p < .05$) festgestellt werden. Für alle anderen Vergleiche zeigten sich keine signifikanten Unterschiede (alle $p \geq .05$).

Es lagen keine signifikanten Wechselwirkungseffekte zwischen dem Messwiederholungsfaktor Zeit und den beiden Kovariaten vor.

In Abbildung 5.5 wird die Leistungsentwicklung der Novizen und der Experten in den Regelaufgaben über die drei Messzeitpunkte hinweg dargestellt. In Tabelle 5.25 erfolgt eine Zusammenfassung der Kovarianzanalyse für die Regelaufgaben.

Die Varianzanalyse mit Messwiederholung zeigte keinen signifikanten Haupteffekt der Expertise. Die Kovariate Visualisierungsfähigkeit trug im Gegensatz zur Kovariate FWG einen statistisch bedeutsamen Anteil zur Aufklärung der Gesamtvarianz bei ($\eta_p^2 = .10$). Um den Effekt der Visualisierungsfähigkeit zu spezifizieren, wurde der korrelative Zusammenhang zwischen der Leistung in den Regelaufgaben und der Leistung im Paper Folding Test berechnet. Die beiden Variablen weisen einen mittelhohen positiven Zusammenhang auf ($r = .47$; $p < .001$). Für den Messwiederholungsfaktor Zeit konnte kein signifikanter Haupteffekt festgestellt werden.

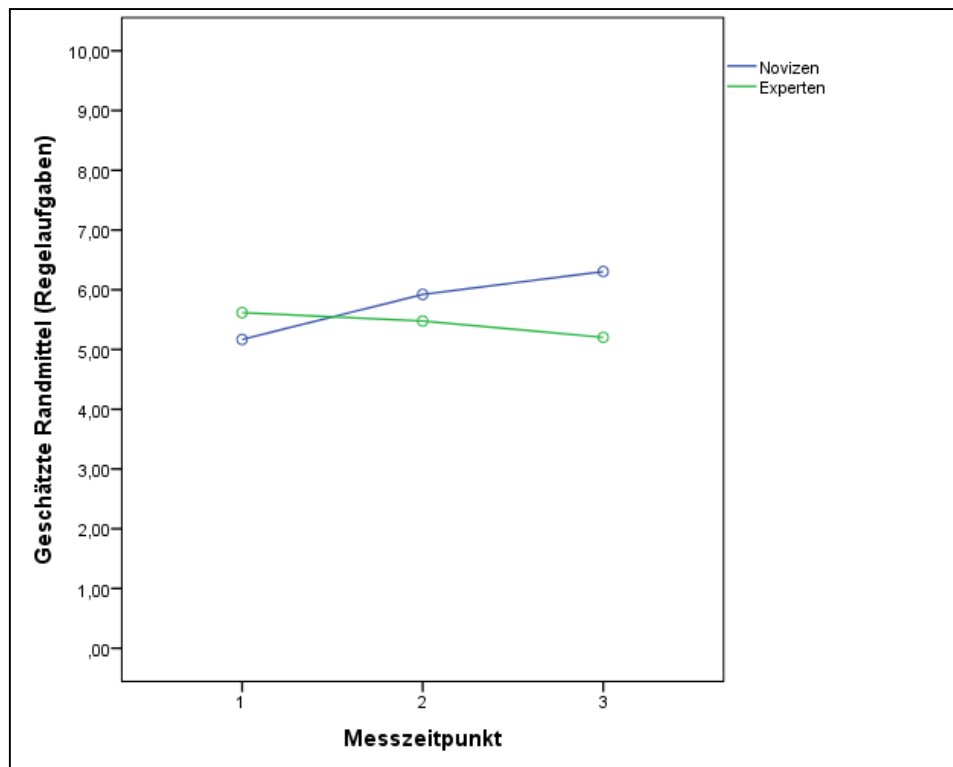


Abbildung 5.5

Leistungsentwicklung der Experten und Novizen in den Regelaufgaben über die drei Messzeitpunkte hinweg (vgl. Malone et al., 2012)

Tabelle 5.25

Ergebnisse der Kovarianzanalyse für die Regelaufgaben

Quelle der Varianz	df	F	p
Gruppenvergleiche			
Expertise (E)	1	1.11	.30
Visualisierungsfähigkeit (V)	1	10.89	.001
FWG (F)	1	.39	.53
Fehler	94	(6.41)	
Messwiederholungsvergleiche ^{a)}			
Zeit (Z)	2	1.03	.35
Z x E	2	5.00	.01
Z x V	2	.69	.49
Z x F	2	1.03	.35
Fehler	170	(2.37)	

Anmerkung. ^{a)}Für den Messwiederholungsfaktor Zeit konnte die Sphäritätsannahme nicht bestätigt werden ($p < .05$). Daher werden nach Greenhouse-Geisser korrigierte Freiheitsgrade und p-Werte berichtet.

Es lag jedoch eine signifikante Interaktion zwischen den Faktoren Zeit und Expertise vor ($\eta_p^2 = .05$). Bei der Betrachtung der Mittelwerte wird deutlich, dass sich die Novizen in den Regelaufgaben über die drei Messzeitpunkte hinweg verbessert haben, während sich die Leistung der Experten nicht verbesserte. Der Scheffé-Test zeigte keine signifikanten Unterschiede über sämtliche Einzelvergleiche (alle $p \geq .05$). Interaktionseffekte zwischen dem Faktor Zeit und den beiden Kovariaten wurden nicht festgestellt.

Zur Bewertung der Leistung in den Reaktionszeitaufgaben wurden, analog zum Vorgehen bei den Experten-Novizen-Vergleichen, zwei abhängige Variablen erhoben: die Anzahl der korrekten Lösungen und die Reaktionszeiten auf die gezeigten Bremsgründe. In Abbildung 5.6 werden die Ergebnisse der Experten und Novizen hinsichtlich der richtigen Lösungen über die drei Messzeitpunkte hinweg dargestellt.

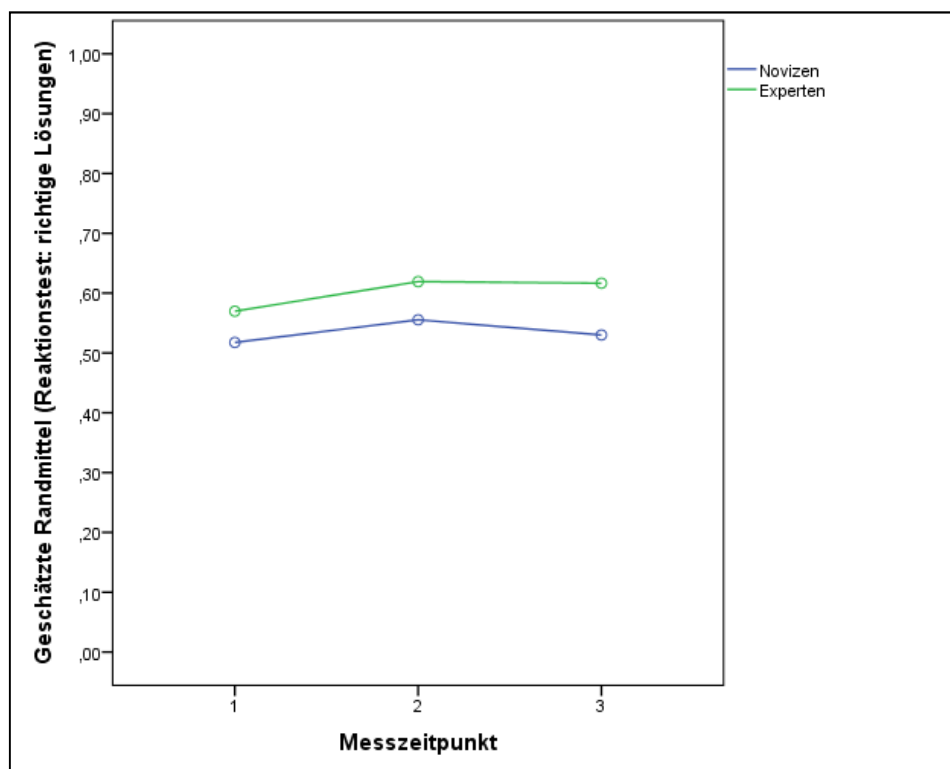


Abbildung 5.6

Leistungsentwicklung der Experten und Novizen in den Reaktionszeitaufgaben (AV: korrekte Lösungen) über die drei Messzeitpunkte hinweg (vgl. Malone et al., 2012)

In Tabelle 5.26 werden die deskriptiven Kennwerte getrennt für die verschiedenen Aufgabentypen aufgeführt.

Tabelle 5.26

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen

Messzeit- punkt		Distraktor		Hazard Perception		Vorschrift		gesamt	
		M	(SD)	M	(SD)	M	(SD)	M	(SD)
1	Novizen	.56	(.24)	.46	(.20)	.50	(.19)	.51	(.21)
	Experten	.58	(.31)	.53	(.18)	.65	(.21)	.59	(.23)
	gesamt	.57	(.27)	.49	(.19)	.57	(.21)	.54	(.22)
	Novizen	.60	(.29)	.53	(.26)	.51	(.23)	.55	(.26)
2	Experten	.71	(.28)	.59	(.21)	.60	(.20)	.63	(.23)
	gesamt	.65	(.29)	.56	(.23)	.55	(.22)	.59	(.25)
	Novizen	.61	(.32)	.48	(.23)	.48	(.26)	.52	(.27)
3	Experten	.73	(.26)	.57	(.20)	.57	(.25)	.62	(.24)
	gesamt	.67	(.29)	.52	(.22)	.52	(.26)	.57	(.26)

Anmerkungen. M = Mittelwert; SD = Standardabweichung

In Tabelle 5.27 erfolgt eine Zusammenfassung der Kovarianzanalyse für die Reaktionszeitaufgaben. Da bei den Reaktionszeitaufgaben zwischen drei Aufgabentypen (Vorschrift, *Hazard Perception* und Distraktor) unterschieden werden konnte, ging in die Kovarianzanalyse neben dem Faktor Zeit als ein weiterer Messwiederholungsfaktor der dreifach gestufte Faktor Aufgabentyp ein.

Die Kovarianzanalyse zeigte einen signifikanten Haupteffekt der Expertise ($\eta_p^2 = .08$): Die Experten erbrachten über die drei Messzeitpunkte hinweg bessere Leistungen im Reaktionstest. Der Faktor Zeit wirkte sich nicht systematisch auf die Leistung im Reaktionstest aus: Es konnten weder ein Haupteffekt noch Interaktionseffekte mit anderen Faktoren oder den Kovariaten festgestellt werden.

Tabelle 5.27

Ergebnisse der Kovarianzanalyse für den Reaktionstest (AV korrekte Lösungen)

Quelle der Varianz	df	F	p
Gruppenvergleiche			
Expertise (E)	1	7.83	.006
Visualisierungsfähigkeit (V)	1	3.56	.06
FWG (F)	1	.19	.67
Fehler	94	(.09)	
Messwiederholungsvergleiche ^{a)}			
Zeit (Z)	2	.99	.37
Z x E	2	.32	.73
Z x V	2	.71	.50
Z x F	2	1.02	.36
Fehler (Z)	188	(.05)	
Aufgabentyp (A)	2	8.31	.001
A x E	2	1.37	.26
A x Z	3	1.63	.18
A x V	2	6.15	.005
A x F	2	1.242	.29
A x E x Z	3	1.05	.37
A x Z x V	3	1.22	.30
A x Z x F	3	.26	.90
Fehler (A)	153	(.10)	
Fehler (A x Z)	305	(.05)	

Anmerkung. ^{a)}Für den Messwiederholungsfaktor Aufgabentyp konnte die Sphäritätsannahme nicht bestätigt werden ($p < .05$). Daher werden nach Greenhouse-Geisser korrigierte Freiheitsgrade und p-Werte berichtet.

Die Kovarianzanalyse deckte allerdings einen Haupteffekt des Messwiederholungsfaktors Aufgabentyp ($\eta_p^2 = .08$) auf. Paarweise Vergleiche (Anpassung des Alpha-Niveaus nach Bonferroni) zeigten, dass die Distraktoren signifikant leichter als die Vorschrift- (mittlere Differenz = .09; $p < .01$) und als die *Hazard-Perception*-Aufgaben (mittlere Differenz = .11; $p < .01$) waren. Vorschrift- und *Hazard-Perception*-Aufgaben unterschieden sich hingegen nicht signifikant in ihrer Schwierigkeit (mittlere Differenz = .03; $p \geq .05$). Es konnte ein tendenzieller Haupteffekt der Visualisierungsfähigkeit festgestellt werden ($\eta_p^2 = .04$), der durch die signifikante Wechselwirkung zwischen dem Faktor

Aufgabentyp und der Kovariate Visualisierungsfähigkeit ($\eta_p^2 = .06$) relativiert wird. Um diese Wechselwirkung zu spezifizieren, wurden Korrelationen der durchschnittlichen Leistungen in den verschiedenen Aufgabentypen mit der Visualisierungsfähigkeit berechnet. Während die Korrelationen von Visualisierungsfähigkeit und Leistung in den Vorschrift- ($r = .47$; $p < .01$) und *Hazard-Perception*-Aufgaben ($r = .23$; $p < .01$) signifikant positiv waren, zeigte sich zwischen der Kovariate und der Leistung bei den Distraktoren kein bedeutsamer Zusammenhang ($r = -.05$; $p = .65$). Weitere statistisch bedeutsame Haupt- und Interaktionseffekte zeigten sich nicht.

Im Folgenden werden die Ergebnisse der Längsschnittuntersuchung für die Reaktionszeiten im Reaktionstest dargestellt. Um zu gewährleisten, dass alle Aufgaben unabhängig von der Länge der Szenarien gleichermaßen in die Analyse eingehen, wurden die Reaktionszeiten z-standardisiert. Die z-Standardisierung erfolgte allerdings auf drei unterschiedliche Arten. Zunächst wurden die einzelnen Reaktionszeiten am Mittelwert der entsprechenden Aufgabe von Experten und Novizen, getrennt für jeden der drei Messzeitpunkte, standardisiert. Der Einbezug der mittleren z-Werte, die auf diese Weise entstehen, erlaubt die Interpretation des Haupteffekts des Faktors Expertise und der Interaktionseffekte zwischen Expertise und Zeit bzw. Aufgabentyp. Eine Interpretation des Haupteffektes des Faktors Zeit ist in diesem Fall allerdings nicht möglich. In Abbildung 5.7 werden die Experten-Novizen-Unterschiede zu jedem der drei Messzeitpunkte dargestellt.

Da negative z-Werte im Zusammenhang mit niedrigeren Reaktionszeiten und damit mit einer besseren Leistung im Reaktionstest stehen, wird aus Abbildung 5.7 deutlich, dass die Experten zu jedem einzelnen Messzeitpunkt im Durchschnitt schneller reagiert haben als die Novizen, der Unterschied zwischen den beiden Gruppen sich allerdings über die Zeit hinweg verringert hat.

In Tabelle 5.28 werden die deskriptiven Kennwerte für die Reaktionszeiten dargestellt. In Tabelle 5.29 werden die interpretierbaren Ergebnisse der Kovarianzanalyse zusammengefasst.

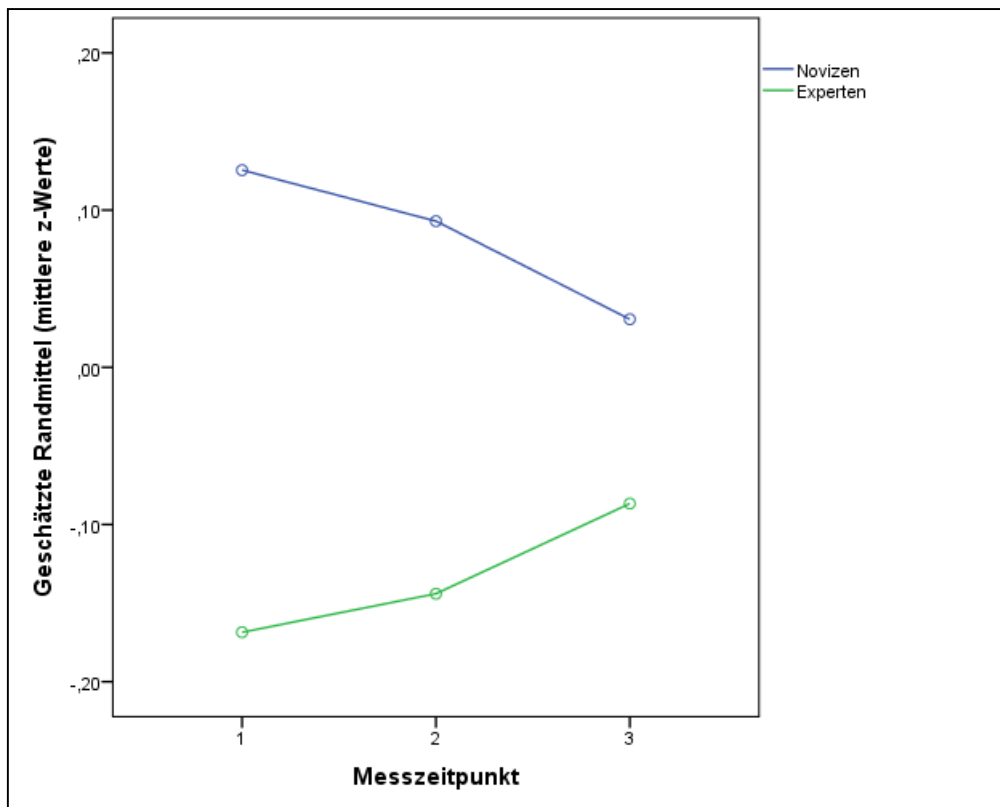


Abbildung 5.7

Leistungsunterschiede zwischen Experten und Novizen in den Reaktionszeitaufgaben (AV: Reaktionszeiten) an drei Messzeitpunkten (vgl. Malone et al., 2012)

Die Kovarianzanalyse zeigte lediglich einen signifikanten Haupteffekt des Faktors Expertise ($\eta_p^2 = .06$). Die z-Werte weisen darauf hin, dass die Experten Bremshinweise innerhalb der Videos früher entdeckt haben als die Novizen. Weitere statistisch bedeutsame Haupt- oder Interaktionseffekte zeigten sich jedoch nicht.

Bei der durchgeführten Kovarianzanalyse kann ein möglicher Haupteffekt der Zeit, aufgrund der z-Standardisierung der Werte der Novizen und Experten zu jedem Messzeitpunkt, nicht interpretiert werden. Daher wurde für eine weitere Analyse die z-Standardisierung am Gesamtmittelwert jeder Aufgabe aller drei Messzeitpunkte gemeinsam durchgeführt. Der Verlauf der mittleren z-Werte, die aus dieser Standardisierung resultierten, ist als Leistungsentwicklung über die Zeit interpretierbar. Die entsprechende Kovarianzanalyse mit dem Mittelwert der berechneten z-Werte als abhängige Variable zeigte keinen generellen Haupteffekt des Faktors Zeit ($F(2, 152) = .69$; $p = .50$). Das heißt die Teilnehmer haben sich über die drei Messzeitpunkte hinweg nicht in statistisch bedeutsamem Maße verbessert oder verschlechtert.

Tabelle 5.28

Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen

Messzeit- punkt		Hazard Perception		Vorschrift		gesamt	
		<i>z</i>	(<i>SD</i>)	<i>z</i>	(<i>SD</i>)	<i>z</i>	(<i>SD</i>)
	Novizen	.11	(.61)	.16	(.65)	.14	(.63)
1	Experten	-.03	(.56)	-.06	(.76)	-.05	(.66)
	gesamt	.04	(.59)	.06	(.70)	.05	(.65)
	Novizen	-.01	(.50)	-.13	(.62)	-.07	(.56)
2	Experten	-.09	(.50)	.20	(.69)	.06	(.60)
	gesamt	-.04	(.50)	.01	(.67)	-.02	(.59)
	Novizen	-.19	(.66)	.07	(.81)	-.06	(.74)
3	Experten	-.01	(.65)	-.01	(.60)	-.01	(.63)
	gesamt	-.11	(.66)	.04	(.72)	-.04	(.70)

Anmerkungen. *M* = Mittelwert; *SD* = Standardabweichung

Die Überprüfung, ob sich eine der beiden Experimentalgruppen in Durchschnitt verbessert oder verschlechtert hat, erforderte die Berechnung einer dritten Art der *z*-Standardisierung. Die Reaktionszeiten bei jeder einzelnen Aufgabe wurden am Gruppenmittelwert der jeweiligen Expertisegruppe standardisiert. Der Mittelwert für eine Aufgabe wurde aus den Reaktionszeiten für dieses Szenario zu allen drei Messzeitpunkten gebildet. Es zeigte sich keine Interaktion zwischen den Faktoren Expertise und Zeit ($F(2, 152) = .69; p = .51$)⁴. Die Experten und die Novizen entwickelten sich hinsichtlich ihrer Geschwindigkeit über die Zeit nicht signifikant verschieden. Paarweise Vergleiche (Anpassung des Alpha-Niveaus nach

⁴ Die beiden Kovarianzanalysen mit den alternativen *z*-Werten werden in Anhang A zusammenfassend dargestellt.

Bonferroni) zeigten, dass für keine der beiden Expertisegruppen signifikante Reaktionszeitunterschiede zwischen den einzelnen Messzeitpunkten vorlagen (alle $p < .05$).

Tabelle 5.29

Ergebnisse der Kovarianzanalyse für den Reaktionstest (AV Reaktionszeiten)

Quelle der Varianz	<i>df</i>	<i>F</i>	<i>p</i>
Gruppenvergleiche			
Expertise (E)	1	4.83	.03
Visualisierungsfähigkeit (V)	1	.13	.72
FWG (F)	1	1.78	.19
Fehler	81	(.87)	
Messwiederholungsvergleiche ^{a)}			
Z x E	2	.95	.39
Aufgabentyp (A)	1	.02	.88
A x E	1	.58	.45
A x Z	2	.41	.66
A x V	1	.13	.72
A x F	1	.00	.99
A x E x Z	2	1.05	.35
A x Z x V	2	.90	.41
A x Z x F	2	.08	.93
Fehler (A)	81	(.30)	
Fehler (A x Z)	152	(.21)	

Die Ergebnisse zeigen, dass die Experten über alle Messzeitpunkte hinweg früher auf die Bremshinweise in den gezeigten Verkehrssituationen reagierten als die Novizen. Weder die Experten noch die Novizen konnten sich darin, wie schnell sie Bremssituationen im Verkehr erkannten, über die Zeit verbessern.

Zusammenfassend hat sich in der längsschnittlichen Betrachtung derjenigen Personen, die an allen drei Messzeitpunkten teilgenommen haben, gezeigt, dass die Experten die Novizen insgesamt nur im Reaktionstest übertroffen haben. Sie lösten über die drei Messzeitpunkte hinweg mehr Reaktionszeitaufgaben richtig und reagierten schneller auf auftauchende Bremshinweise. In den Wissensaufgaben und Regelaufgaben zeigte sich eine Interaktion zwischen den Faktoren Expertise und Zeit: Insbesondere in den Wissensaufgaben verbesserten sich nur die Novizen über die Zeit. Die Leistung in den Regelaufgaben zeigte

einen deutlichen positiven Zusammenhang mit der Kovariate Visualisierungsfähigkeit. Diese Kovariate wies in Wechselwirkung mit dem Faktor Aufgabentyp einen deutlichen Zusammenhang mit der Leistung im Reaktionstest (AV korrekte Lösungen) auf. Es zeigte sich zudem, dass die Distraktoren weniger schwierig waren als die anderen Aufgabentypen. Während keinerlei Zusammenhänge zwischen der Visualisierungsfähigkeit und der Leistung in den Distraktoren bestand, korrelierte diese Kovariate deutlich positiv mit der Leistung in den beiden anderen Aufgabentypen. Weder die Experten noch die Novizen konnten ihre Leistungen im Reaktionstest über die Zeit steigern.

5.4 Diskussion der Längsschnittstudie

Die Ergebnisse der Längsschnittstudie werden im Folgenden hinsichtlich ihrer Bedeutung für die Bestätigung der aufgestellten Hypothesen diskutiert und mögliche Einschränkungen der Studie dargestellt.

5.4.1 Bestätigung der Hypothesen in der Längsschnittstudie

5.4.1.1 Messzeitpunkt 1

Es wurde vermutet, dass die Experten die Novizen in den Aufgaben, die Verkehrswissen und Faustregeln abprüfen, zum ersten Messzeitpunkt deutlich übertreffen (Hypothese 1). Allerdings zeigte sich im Experten-Novizen-Vergleich, dass zum Messzeitpunkt 1 keine deutlichen Leistungsunterschiede zwischen den Experten und den Novizen in den Wissensaufgaben bestanden. Die Hypothese wurde demzufolge durch die Ergebnisse dieser Studie nicht gestützt. Dies deckt sich mit den Ergebnissen der ersten Vorstudie, in der ebenfalls keine Expertiseunterschiede festgestellt werden konnten. Die Ergebnisse stehen im Kontrast zu den Befunden der Expertiseforschung in anderen Domänen, die besagen, dass Experten sich auf ihrem Fachgebiet durch ein umfangreiches Wissen auszeichnen (Gruber & Mandl, 1995). Das Abprüfen deklarativer Wissensinhalte im Bereich der Gefahrenlehre scheint demnach keine geeignete Methode zu sein, Fahrexpertise valide zu messen. Zwischen erfahrenen Fahrern und Fahrschülern, die sich noch ganz am Beginn ihrer Fahrausbildung befinden, sollten deutliche Unterschiede beim Fahrenkönnen bestehen, die sich durch die Wissensaufgaben allerdings nicht abbilden lassen. Die Tatsache, dass die Experten nicht deutlich mehr Wissensaufgaben gelöst haben als die Novizen, weist darauf hin, dass Aufgaben zu deklarativen Wissensinhalten zum Thema Gefahrenlehre nicht geeignet sind um zwischen Experten und Novizen zu differenzieren. Zum Teil müssen diese Inhalte vermutlich

gar nicht explizit erworben werden, da sie sich durch eher allgemeine Erfahrungen erschließen lassen (Aufgaben, die sowohl von Experten als auch von Novizen richtig gelöst werden). Andere Wissensinhalte erscheinen bei der Expertiseentwicklung und zum sicheren Fahren irrelevant (Aufgaben, die sowohl von Experten als auch von Novizen nicht richtig gelöst werden).

Die Annahme, dass die Experten den Novizen in den Regelaufgaben zum ersten Messzeitpunkt überlegen sein sollten, wurde nur tendenziell durch die Ergebnisse des Experten-Novizen-Vergleichs der Längsschnittstudie bestätigt. Befunde der Expertiseforschung, die in der Regel deutliche Unterschiede zwischen dem Umfang des domänenspezifischen Wissens der Experten und der Novizen aufzeigen, wurden in dieser Studie demnach nicht repliziert. Das Wissen um die im Fahrschulunterricht erlernten Faustregeln scheint somit nicht spezifisch für die Domäne Autofahren zu sein. Es kann angenommen werden, dass Experten die Regeln nicht explizit beim Fahren einsetzen.

In Anlehnung an die Ergebnisse der Forschung zu *Hazard Perception* (vgl. Kapitel 2.3.3) und der eigenen Ergebnisse mit neu entwickelten Reaktionszeitaufgaben zur Erfassung der Reaktion auf Bremshinweise in Verkehrsszenarien (Vorstudie 2, vgl. Kapitel 4.2) wurde angenommen, dass die Experten am ersten Messzeitpunkt zu einer zuverlässigeren Bremshinweise in dargebotenen Verkehrsszenarien entdecken (Hypothese 3) und zum anderen auch schneller darauf reagieren als die Novizen (Hypothese 4). In Übereinstimmung mit den internationalen Forschungsergebnissen konnte gezeigt werden, dass die Experten mehr Bremshinweise in den Verkehrsvideos entdeckten als die Novizen (z.B. McKenna & Crick, 1994a; McKenna & Crick, 1994b). Hypothese 3 wird damit durch die Ergebnisse der Studie gestützt. Außerdem reagierten die Experten zu Messzeitpunkt 1 früher auf in den Videos auftauchende Bremshinweise, als die Novizen. Dieser Befund stützt somit Hypothese 4. Insgesamt konnten demnach die Ergebnisse von Vorstudie 2 für den ersten Messzeitpunkt der Längsschnittstudie repliziert werden. Dies unterstützt die Annahme, dass es sich bei der Gefahrenwahrnehmung um eine expertiserelevante Fähigkeit handelt. Außerdem wurde erneut bestätigt, dass der entwickelte Reaktionstest als ein kriteriumsvalides Messinstrument für Fahrexpertise gelten kann.

5.4.1.2 Messzeitpunkte 2 und 3

In der Zeit zwischen dem Beginn der Studie und dem Ende der Studie, sollten die Novizen theoretisches Wissen und etwas praktische Übung gesammelt haben. Die Fähigkeit zum Wissenstransfer wird erst durch Erfahrung ausgebildet (z.B. Dreyfus, S. E. & Dreyfus, 1980;

Spiro et al., 1988). Da die Novizen bis zum dritten Messzeitpunkt nur vergleichsweise wenig Erfahrung im selbstständigen Fahren gesammelt haben sollten, wurde angenommen, dass die Experten den Novizen auch zu späteren Messzeitpunkten im Reaktionstest, der den flexiblen Einsatz des Wissens verlangt, überlegen sein sollten. Diese Annahme wurde aus der Überlegung abgeleitet, dass es sich bei den Reaktionszeitaufgaben um handlungsnahen Aufgabenformate handelt, die ähnliche Anforderungen wie das Fahren selbst an den Teilnehmer stellen und die folglich nur durch Handlung (hier durch tatsächliches Fahren) trainiert werden können. Dabei sollten die Experten zu den Messzeitpunkten 2 und 3 zum einen mehr Aufgaben richtig lösen (Hypothese 5) und zum anderen Bremshinweise schneller entdecken als die Novizen (Hypothese 6). Es zeigte sich, dass die Experten den Novizen zu Messzeitpunkt 2 hinsichtlich der Anzahl der richtig gelösten Aufgaben nicht überlegen waren und zu Messzeitpunkt 3 nur tendenziell. Zudem reagierten die Experten zu Messzeitpunkt 2 und 3 nur tendenziell früher auf die gezeigten Bremshinweise. Die längsschnittliche Betrachtung derjenigen Personen, die an allen drei Messzeitpunkten anwesend waren, zeigte für beide abhängigen Variablen einen Haupteffekt des Faktors Expertise. Über alle drei Messzeitpunkte hinweg lösten die Experten mehr Aufgaben richtig und reagierten schneller. Die Ergebnisse stützen die beiden Hypothesen zum größten Teil. Es kann davon ausgegangen werden, dass die Novizen am Ende ihrer Ausbildung das Leistungsniveau der Experten im Reaktionstest noch nicht erreicht haben. Der Expertiseunterschied, der erwiesenermaßen nach der formellen Fahrausbildung noch immer zwischen unerfahrenen und erfahrenen Fahrern besteht, kann demzufolge mit Hilfe des Reaktionstests abgebildet werden. Die Kriteriumsvalidität der Reaktionszeitaufgaben ist als hoch einzuschätzen, da es mithilfe dieses Aufgabenformates auch noch möglich war Novizen, die ihre Fahrausbildung bereits abgeschlossen hatten, von Experten zu trennen.

5.4.1.3 Längsschnittstudie

Bis zu den beiden Messzeitpunkten 2 und 3 hatten die Novizen bereits zum größten Teil viele theoretische und praktische Fahrstunden absolviert, beziehungsweise hatten ihre Fahrerlaubnis bereits erworben und etwas Erfahrung im selbstständigen Fahren gesammelt. Die Experten hingegen sollten sich innerhalb der Zeit der Längsschnittstudie kaum neues verkehrsspezifisches Wissen aneignen. Da innerhalb der Fahrausbildung das notwendige Wissen erworben werden sollte, um die Aufgaben zum Wissen bezüglich Gefahrenlehre und Faustregeln erfolgreich zu lösen, wurde angenommen, dass die Novizen sich in den Wissensaufgaben (Hypothese 7) und in den Regelaufgaben (Hypothese 8) über die Zeit

deutlicher verbessern als die Experten, die in der Zeit keine aktiven Bemühungen verfolgen sollten, ihr Wissen um Gefahren im Verkehr und um Regeln zu erweitern.

Es zeigte sich für die Leistung in den Wissensaufgaben eine signifikante Wechselwirkung zwischen den beiden Faktoren Expertise und Zeit. Die Novizen steigerten ihre Leistung über die Zeit, während die Leistung der Experten konstant blieb. Die Ergebnisse stützen somit Hypothese 7. Zwischen dem zweiten und dem dritten Messzeitpunkt zeigte sich allerdings auch bei den Novizen kein bedeutsamer Leistungsunterschied. Dieses Ergebnis erscheint plausibel, da viele Novizen bereits zum zweiten Messzeitpunkt entweder ihre theoretische oder bereits die gesamte Fahrausbildung abgeschlossen hatten. Ein weiterer Zuwachs an deklarativen Wissensinhalten, die in der Regel Inhalt des theoretischen Fahrschulunterrichtes sind, war daher nicht zu erwarten. Da die Wissensaufgaben den Wissenszuwachs der Novizen während der Fahrschulbildung abbilden, sind sie als lehrzielvalide zu bezeichnen.

Für die Leistung in den Regelaufgaben zeigte sich in der Längsschnittstudie ebenfalls eine signifikante Wechselwirkung der Faktoren Expertise und Zeit. Hypothese 8 wird daher durch die Ergebnisse der Studie gestützt. Die deskriptive Statistik weist darauf hin, dass die Novizen über die Zeit mehr Regelaufgaben lösen, während die Experten sich weder verbessern noch verschlechtern. Dass die Novizen sich in den Regelaufgaben nicht so deutlich verbessert haben wie in den Wissensaufgaben, spiegelt die Tatsache wider, dass die Regelaufgaben im Gegensatz zu den Wissensaufgaben in dieser Studie nicht so stark mit den Aufgaben der tatsächlichen Fahrerlaubnisprüfung übereinstimmen. Für die Regelaufgaben konnte somit nicht explizit gelernt werden. Dennoch können sie zumindest als tendenziell lehrzielvalide betrachtet werden, da die angenommene Wechselwirkung nachgewiesen werden konnte.

Der Befund, dass es zu keiner generellen Verbesserung der Teilnehmer in den Wissens- und Regelaufgaben gekommen ist, zeigt, dass es sich bei dem gefundenen Lerneffekt der Novizen, nicht um einen Trainingseffekt handelt, der durch die dreimalige Bearbeitung der Wissensaufgaben innerhalb der Studie selbst zustande gekommen ist.

Auch für die Reaktionszeitaufgaben wurde angenommen, dass die Novizen ihre Leistungen über die Zeit deutlicher steigern als die Experten. Zum einen sollten sie immer mehr Bremshinweise innerhalb der gezeigten dynamischen Darstellungen von Verkehrsszenarien richtig identifizieren (Hypothese 9) und zum anderen immer schneller darauf reagieren (Hypothese 10) während die Experten eher ein gleichbleibendes

Leistungsniveau aufweisen sollten. Es zeigten sich weder für die abhängige Variable korrekte Lösungen noch für die Reaktionszeiten in der Längsschnittstudie signifikante Wechselwirkungen zwischen Zeit und Expertise. Demnach wurden Hypothese 9 und 10 nicht durch die Ergebnisse gestützt. Die beiden Gruppen hatten also ein unterschiedliches Leistungsniveau – die Experten waren überlegen – ihre Leistungsentwicklung war jedoch gleich. Weder die Experten noch die Novizen verbesserten sich über die Zeit, was gegen einen generellen Trainingseffekt durch die mehrmalige Darbietung der Aufgaben spricht. Dass sich die Novizen nicht während ihrer Fahrschulausbildung und kurz danach verbesserten, spricht dafür, dass die zur erfolgreichen Bewältigung der Reaktionszeitaufgaben erforderlichen Fähigkeiten in dieser Zeit nicht erworben wurden. Zumindest erfolgte dies nicht in mithilfe des eingesetzten Verfahrens messbarem Ausmaß. Dass die Experten ihre Leistungen nicht deutlich steigern konnten, ist darauf zurück zu führen, dass sie vermutlich ein bereits hohes Niveau innerhalb eines halben Jahres nicht zusätzlich steigern konnten. Diese Erklärung lässt sich durch die Beobachtung stützen, dass in den ersten drei Jahren nach dem Führerscheinerwerb eine exponentielle Verringerung des Unfallrisikos festgestellt werden kann (vgl. Schade, 2001). Danach besteht nur ein Restrisiko, das sich kaum weiter verringert. Wenn man umgekehrt annimmt, dass dem Rückgang des Unfallrisikos ein exponentielles Wachstum an relevanten Fähigkeiten zu Grunde liegt, erklärt dies, warum eine Leistungssteigerung der Experten innerhalb eines halben Jahres nicht messbar war.

Der Tatsache, dass sich die Novizen nicht in den Reaktionszeitaufgaben verbesserten, könnten verschiedene Aspekte zugrunde liegen. Einerseits haben nicht alle Novizen die Fahrerlaubnis innerhalb der Zeit der Längsschnittstudie erworben. Nur 30 der 50 Novizen gaben an, am dritten Messzeitpunkt bereits selbstständig gefahren zu sein. Die übrigen Personen befanden sich noch in der Fahrausbildung. In der derzeitigen Fahrausbildung wird die schnelle korrekte Reaktion auf Verkehrsgegebenheiten weder im Realverkehr noch innerhalb von Computersimulationen explizit trainiert. Dies lässt die Folgerung zu, dass Personen innerhalb der Fahrausbildung keine Fertigkeiten zur schnellen Reaktion auf Verkehrsgegebenheiten erwerben. Selbst die Personen, die bis zum letzten Messzeitpunkt eine Fahrerlaubnis erworben haben, wiesen erst eine geringe Fahrleistung auf (90 % maximal 2000 Kilometer). Es kann angenommen werden, dass das Instrument nicht hinreichend sensitiv war, um sehr kleine Fortschritte zu messen. Das Verfahren ist damit zur Erfassung der Leistungsentwicklung innerhalb der Fahrausbildung nicht inhaltsvalide. Dies bedeutet aber

zudem, dass in der Ausbildung Aspekte nicht trainiert werden, die die spätere Expertise eines Fahrers tatsächlich ausmachen.

Insgesamt zeigte die Längsschnittstudie, dass die Novizen im Laufe ihrer Fahrausbildung Wissen um Gefahren im Verkehr und Faustregeln erwerben, die Zeit allerdings nicht auszureichen scheint, um relevante Reize im Verkehr immer schneller zu entdecken. Die Novizen erreichen innerhalb ihrer Fahrausbildung in denjenigen Aufgaben, die deklarative Wissensinhalte explizit abfragen, schnell das Niveau der erfahrenen Fahrer. Im handlungsnahen Aufgabenformat, das es verlangt, Verkehrssituationen schnell zu überblicken und zu entscheiden, ob eine Reaktion notwendig ist, erreichen sie das Niveau der erfahrenen Fahrer jedoch nicht.

Hinsichtlich des Einflusses des Aufgabentyps bei den Reaktionszeitaufgaben sollten durch die Längsschnittstudie Hinweise darauf gefunden werden, wie sich die Leistung in den verschiedenen Aufgabentypen entwickelt (Forschungsfrage 1). Es zeigte sich, dass die Aufgabentypen insgesamt unterschiedlich schwierig waren. Über alle Messzeitpunkte hinweg waren die Distraktoren weniger schwierig als die *Hazard-Perception*- und Vorschriftaufgaben. Ein Schwierigkeitsunterschied zwischen den *Hazard-Perception*- und den Vorschriftaufgaben konnte im Gegensatz zu Vorstudie 2 nicht festgestellt werden. Zum ersten Messzeitpunkt reagierten die Teilnehmer schneller auf die salienten Bremshinweise in den *Hazard-Perception*-Aufgaben als auf die impliziten Bremshinweise in den Vorschriftaufgaben. Wechselwirkungen hinsichtlich Zeit und Expertise lagen jedoch nicht vor.

5.4.2 Zusammenhang von Leistung und räumlichen Fähigkeiten

Zur Ermittlung der Zusammenhänge zwischen verkehrsspezifischen Fähigkeiten und räumlichen Fähigkeiten wurden den Teilnehmern zwei Tests zu den räumlichen Fähigkeiten Visualisierungsfähigkeit und Fähigkeit zur Wahrnehmung von Geschlossenheit vorgelegt. Es sollte ermittelt werden, inwiefern Zusammenhänge zwischen den Leistungen in den beiden Tests und der Leistung in den verschiedenen verkehrsspezifischen Leistungsmaßen vorliegen (Forschungsfrage 2), und ob sich die Höhe des Zusammenhangs über die Zeit verändert (Forschungsfrage 3). Zusammenfassend konnte gezeigt werden, dass bei den Experten räumliche Fähigkeiten und die Leistung in den entwickelten verkehrsspezifischen Aufgabenformaten einen deutlicheren positiven Zusammenhang aufweisen. Konsistent für beide Gruppen zeigte sich, dass hohe räumliche Fähigkeiten mit guten Leistungen (viele richtig beantwortete Aufgaben, niedrige Reaktionszeiten) zumindest tendenziell einhergingen.

Die Kovarianzanalyse zu den verschiedenen Aufgabenformaten deckte auf, dass die Kovariaten teilweise einen bedeutsamen Anteil der Varianz zusätzlich zu den anderen Faktoren erklärten. Besonders deutlich war der Zusammenhang zwischen der Visualisierungsfähigkeit und der Leistung in den Regelaufgaben. Dieser Befund stützt die Annahmen der *ability-as-compensator*-Hypothese, nach der Personen mit guten räumlichen Fähigkeiten die schlechte Darstellungsqualität von Arbeitsmaterial ausgleichen können (Mayer & Sims, 1994). Bei den Regelaufgaben handelte es sich um den einzigen Aufgabentyp, bei dem kein Bildmaterial zur Veranschaulichung gezeigt wurde; Dieses Aufgabenformat kann daher als von geringer Darstellungsqualität betrachtet werden. Möglicherweise können sich Personen, die über eine gute Visualisierungsfähigkeit verfügen, die in den Regelaufgaben verbalisiert dargestellten Verkehrsszenarien (Abstände und Geschwindigkeiten) vorstellen und dadurch die fehlende bildhafte Darstellung kompensieren. Es ist allerdings zu beachten, dass Kausalerklärungen auf der Grundlage korrelativer Zusammenhänge nicht möglich sind. Denkbar ist daher auch, dass sich die beiden Fähigkeiten nicht gegenseitig bedingen, sondern beide von einer dritten Variablen, wie z.B. Motivation, in ähnlicher Weise beeinflusst werden.

Die Fähigkeit zur Wahrnehmung von Geschlossenheit stand in einem deutlicheren Zusammenhang zu der Leistung im Reaktionstest. Zum einen kann dies durch die Zeitkomponente erklärt werden, die sowohl für die Leistung im Hidden Patterns Test als auch im verkehrsspezifischen Reaktionstest von Bedeutung ist. Zum anderen kann man annehmen, dass Personen, die generell über die Fähigkeit verfügen, Figuren in einem komplexen Gefüge zu erkennen, auch Vorteile bei der Identifikation relevanter Reize im komplexen Verkehrsgeschehen haben.

Wechselwirkungen zwischen dem Faktor Zeit und den räumlichen Kovariaten wurden für keine der verkehrsspezifischen Aufgabenarten gezeigt. Dies lässt darauf schließen dass sich die Zusammenhänge zwischen den Kovariaten und den Leistungen in den verschiedenen Aufgabenformaten nicht über die Zeit verändern.

6 Diskussion und Ausblick

Dieses Kapitel beginnt mit der Darstellung des Beitrags der Ergebnisse der vorgestellten Studien zum Verstehen der Beschaffenheit und des Erwerbs von Expertise in der Domäne Autofahren. Dazu werden die Ergebnisse im Zusammenhang mit den generellen Befunden zur Expertiseforschung über verschiedene Domänen hinweg betrachtet. Außerdem werden die Befunde der vorliegenden Arbeit den Ergebnissen einschlägiger Studien der verkehrspsychologischen Forschung gegenüber gestellt. Zudem wird diskutiert, inwieweit durch die Ergebnisse der vorgestellten Studien aber auch infolge ihrer methodischen Grenzen neue Forschungsdesiderate entstanden sind und wie weitere Forschungsvorhaben in diesem Zusammenhang gestaltet sein sollten. Zum Schluss wird die praktische Bedeutung der Ergebnisse für die Optimierung der Fahrerlaubnisprüfung und für die Fahrausbildung dargestellt.

6.1 Expertise und Expertiseerwerb in der Domäne Autofahren

Die in dieser Arbeit vorgestellten Studien wurden mit dem Ziel durchgeführt, Einblicke in die Besonderheiten einerseits des Charakters der Fahr- und Verkehrsexpertise und andererseits der Entwicklung von Expertise in dieser Domäne zu erlangen. Im Folgenden soll zunächst diskutiert werden, inwiefern Fahrexpertise in ihrer Beschaffenheit Ähnlichkeiten zur Expertise in anderen Domänen aufweist und anschließend welche Modelle des Expertiseerwerbs zur Erklärung der Entwicklung von Expertise in der Domäne Autofahren geeignet sind.

6.1.1 Merkmale von Fahrexpertise

Gemäß dem ersten Stadium des in Kapitel 2.1.4 vorgestellten *expert-performance*-Ansatzes von Ericsson und Smith (1991a) wurden in den beiden ersten Vorstudien Aufgaben entwickelt und getestet, die Expertise in der Domäne Autofahren erfassen sollten. Die durchgeführten Experten-Novizen-Vergleiche erlauben Aussagen darüber, was die Experten in der Domäne Autofahren gegenüber den Novizen auszeichnet. Eine allgemein akzeptierte Definition von Expertise besagt, dass sie sich in einer stabilen hohen Leistungsfähigkeit in domänenspezifischen Aufgaben zeigt (Gruber, 1994; Posner, 1988). Wenn man dieser Hypothese als einer generellen Aussage vertraut, kann man im Umkehrschluss davon ausgehen, dass solche Aufgaben, in denen Experten die Novizen nicht übertreffen, keine Anforderungen stellen, die für die Domäne Autofahren relevant sind. Aus den Ergebnissen

der ersten Vorstudie und der Längsschnittstudie dieser Arbeit könnte daher abgeleitet werden, dass es sich bei deklarativem Verkehrswissen nicht um ein Merkmal von Fahrexpertise handelt. Die Experten zeigten keine besseren Leistungen als die Novizen, wenn verkehrsspezifisches Wissen abgefragt wurde. Bei den Regelaufgaben, die das Wissen um die sogenannten Faustregeln zur Einschätzung von Anhalteweg, Sicherheitsabstand und Überholweg abprüfen, zeigte sich in der Längsschnittstudie ein ähnliches Bild. Diese Ergebnisse scheinen nicht nur zu den Befunden vieler anderer Studien aus unterschiedlichen Domänen (z.B. Bédard & Chi, 1992; Posner, 1988), sondern auch zur Definition von Expertise selbst im Gegensatz zu stehen (vgl. Gruber & Mandl, 1995). Eine enorme Basis domänenspezifischen Wissens wurde besonders konsistent als Merkmal von Expertise identifiziert (Bédard & Chi, 1992). Laut den Autoren kann die Übertragbarkeit dieses Ergebnisses, der zumeist auf Laborstudien basierenden Expertiseforschung, auf Aspekte des alltäglichen Lebens – also ihre ökologische Validität – jedoch angezweifelt werden. Die Interpretation, spezifische deklarative Wissensinhalte seien für sicheres Fahren völlig irrelevant, ist dennoch anzuzweifeln. Das Abprüfen dieser Wissensinhalte erfolgte in den Studien dieser Arbeit durch Aufgaben im Multiple-Choice-Antwortformat. Dies erforderte die Repräsentation der abzu prüfenden Kenntnisse auf einem bewussten Niveau, da wohlüberlegte Entscheidungen darüber getroffen werden müssen, welche der vorgegebenen Antwortalternativen falsch bzw. richtig sind. Möglicherweise stellt dies jedoch aufgrund der geringen ökologischen Validität der Aufgabenstellung eine unangemessene Form der Überprüfung des Expertenwissens dar. Die Ergebnisse der Vorstudie 2 der vorliegenden Arbeit stützen diese Annahme. Die Experten lösten in einem Reaktionstest, in dem auf Bremshinweise in gezeigten Videos reagiert werden musste, mehr Aufgaben richtig als die Novizen. Dieser Unterschied zeigte sich ebenfalls, wenn man nur solche Aufgaben betrachtete, in denen verkehrsspezifisches Wissen relevant zur Identifikation der Bremshinweise war. Dieses Aufgabenformat setzt zwar voraus, dass das erforderliche Wissen verfügbar ist, aber nicht, dass es bewusst verarbeitet wird. Es wird viel eher auf einem anwendungsbezogenen Niveau abgeprüft. Man kann demnach annehmen, dass verkehrsspezifisches Wissen erst als expertiserelevant aufgefasst werden kann, wenn es auch adäquat eingesetzt werden kann. Wie in der Längsschnittstudie dieser Arbeit deutlich wurde, weisen Fahranfänger und erfahrene Fahrer eine ähnlich umfangreiche Wissensbasis auf. Der Leistungsunterschied der beiden Gruppen zeigte sich erst, wenn die situationsgerechte Anwendung des Wissens in simulierten Verkehrssituationen erforderlich wurde, wie es im

Reaktionstest der Fall war. Dies spricht für einen qualitativen Unterschied zwischen dem Experten- und dem Novizenwissen in der Domäne Autofahren, der es erforderlich macht, zur Leistungsüberprüfung geeignete Methoden einzusetzen.

Gestützt wird diese Erklärung durch einen weiteren allgemeinen Befund der Expertiseforschung, der besagt, dass die hohe Leistungsfähigkeit der Experten domänenspezifisch ist. Das heißt, Experten haben gegenüber Novizen nur dann Vorteile, wenn die Aufgaben, die sie zu erfüllen haben, spezifisch für die jeweilige Domäne sind, also je ähnlicher die Anforderungen der Aufgaben den Anforderungen sind, die in der Domäne normalerweise gestellt werden. Demzufolge zeigte sich der Unterschied zwischen den Experten und den Novizen in den vorgestellten Studien deutlicher, wenn Aufgaben im Reaktionszeitformat und nicht im Multiple-Choice-Aufgabenformat eingesetzt wurden, weil die Reaktionszeitaufgaben spezifischer für die Domäne Autofahren sind. Bereits De Groot (1965) führte an, dass Aufgaben, die besser den Anforderungen des natürlichen Verhaltens entsprechen, bessere Prädiktoren für Leistung seien als die von ihm eingesetzten Gedächtnisaufgaben. In Experten-Novizen-Vergleichen in verschiedenen Sportarten zeigten sich umso höhere Effektstärken, je ökologisch valider die Studien waren (Thomas et al., 2003). Methodenvergleichende Studien, in denen der Faktor *Handlungsnähe* des Aufgabenformates experimentell modifiziert wurde, sind jedoch bisher nicht bekannt. Zur Analyse, inwieweit dieser Faktor den Leistungsunterschied zwischen Experten und Novizen beeinflusst, ist die Durchführung weiterer Studien wünschenswert. Dabei sollte der Faktor Handlungsnähe bei identischem Stimulusmaterial (Verkehrsszenarien) experimentell variiert werden. Dadurch kann mithilfe von Experten-Novizen-Vergleichen die Annahme überprüft werden, dass mit steigender Handlungsnähe des Antwortformates die Fähigkeit zur Trennung von Könnern und Nicht-Könnern der Aufgaben ansteigt. Es kann vermutet werden, dass sich als Ergebnis eine Wechselwirkung zwischen dem Faktor Expertise und dem Faktor Handlungsnähe zeigt. Wenn diese Wechselwirkung so geartet ist, dass die Überlegenheit erfahrener Fahrer gegenüber Fahrnovizen mit steigender Handlungsnähe des Antwortformates deutlicher ausfällt, spricht dies für einen validitätssteigernden Effekt des Faktors Handlungsnähe.

In Übereinstimmung zu den Befunden der Forschung zur Gefahrenwahrnehmung (vgl. Biermann, 2007; McKenna & Crick, 1994b; Smith et al., 2009a; Sümer et al., 2007) zeigte sich, dass die Experten Gefahrenhinweise innerhalb der präsentierten Verkehrsszenarien sicherer und schneller erkannten. Dies unterstützt die Annahme, dass die

Gefahrenwahrnehmung eine sicherheitsrelevante Fähigkeit ist, die durch Reaktionstests valide überprüft werden kann. Innerhalb der vorliegenden Arbeit konnte zudem nachgewiesen werden, dass dies nicht nur für die Entdeckung von tatsächlichen Gefahren im Verkehr gilt, sondern besonders auch für die Identifikation relevanter Verkehrsgegebenheiten, die den Fahrer auf vorschriftsgemäßes Verhalten hinweisen. Durch ihre hohe Auftretenswahrscheinlichkeit beim alltäglichen Fahren, sollte die Umsetzung von Verkehrsregeln bei erfahrenen Fahrern besonders gut beherrscht werden.

Die Ähnlichkeit zur Expertise im Sport, die sich darin zeigt, dass die Experten, verglichen mit Novizen, akkuratere und schnellere Entscheidungen treffen (Hodges et al., 2007), wurde in Vorstudie 2 und auch in der Längsschnittstudie für die Fähigkeit zum Überblicken komplexer Verkehrssituationen gezeigt. Offen bleibt dabei allerdings, welche Komponenten der Entscheidungsfindung bei der Entstehung des Leistungsunterschiedes zwischen Experten und Novizen tatsächlich eine Rolle spielen. Die Frage, *warum* ein Novize auf einen relevanten Hinweisreiz nicht reagiert hat, bleibt unbeantwortet. Diese Unsicherheit liegt darin begründet, dass innerhalb der vorliegenden Arbeit nur das erste Stadium des *expert performance approach* (Ericsson & Smith, 1991a) erfolgreich absolviert wurde: Es wurden repräsentative Aufgaben entwickelt, die sich im Experten-Novizen-Vergleich als valide erwiesen haben. Zur genauen Darstellung der Expertise in einer Domäne fehlt allerdings das zweite Stadium im Forschungsprozess. Innerhalb der vorgestellten Studien wurde nicht untersucht, welche Prozesse und Mechanismen der Leistung der Experten in den repräsentativen Aufgaben zugrunde liegen. Weitere Forschung im Bereich der Erfassung der Wahrnehmung und Verarbeitung von Verkehrsgegebenheiten, als ein Merkmal von Expertise in der Domäne Autofahren, sollte sich demnach mit der Analyse der kognitiven Prozesse befassen, die der Bearbeitung der Aufgaben (z.B. dem Treffen einer Entscheidung im Reaktionstest) zugrunde liegen.

Einen Anhaltspunkt kann dabei die Analyse von Blickbewegungen bei der Bearbeitung der Aufgaben liefern. Experten-Novizen-Vergleiche ermöglichen dabei Hinweise darauf, inwiefern erfahrene und unerfahrene Fahrer sich in qualitativen und quantitativen Aspekten der Beobachtung von Verkehrsszenarien unterscheiden. Dadurch könnte erklärt werden, welche Beobachtungsprozesse einer Entscheidung bezüglich der Präsenz von relevanten Hinweisreizen innerhalb eines Szenarios vorangehen. Blickbewegungsmessungen bei Simulationen zur Leistungsmessung im Sport haben gezeigt, dass das Blickverhalten von Experten sich von dem der Novizen unterscheidet und angepasst an die jeweiligen

Anforderungen variiert (Ward, Williams & Hancock, 2006). Forschung in der Domäne Verkehr hat offenbart, dass effektives Absuchen der Umwelt mit besseren Gefahrenwahrnehmungsleistungen einhergeht (z.B. Crundall et al., 1998; Mourant & Rockwell, 1972). Für die Prozesse bei der Exploration statischer Präsentationen von Verkehrsszenarien konnten Huestegge, Skottke, Anders, Muesseler und Debus (2010) jedoch anhand von Blickbewegungsdaten zeigen, dass sich Experten und Novizen nicht darin unterscheiden, wie früh sie relevante Hinweisreize ansahen, sondern wie schnell ihre Reaktion darauf erfolgte. Das bedeutet, die schnelle Reaktionszeit der Experten sollte darin begründet liegen, dass Experten eine schnellere Bewertung der Reize vornehmen als die Novizen. Ein Nachweis dafür, dass diese Ergebnisse sich auch auf die Verarbeitung dynamisch präsentierter Verkehrsszenarien anwenden lassen, steht bisher noch aus und stellt damit ein Forschungsdesiderat für die zweite Stufe im Prozess der Expertiseforschung in der Domäne Autofahren dar. Dadurch könnte beispielsweise ermittelt werden, inwiefern möglicherweise frühes antizipatorisches Blickverhalten eine schnelle Bewertung der Situation und damit die Entscheidung begünstigt.

Ein weiteres Merkmal auf der Stufe der Analyse kognitiver Mechanismen bei der Verarbeitung von Aufgaben zur Wahrnehmung kann die Erfassung der kognitiven Belastung darstellen. Es hat sich als ein Merkmal von Expertise gezeigt, dass die kognitive Belastung bei der Ausübung einer Tätigkeit mit steigender Fähigkeit sinkt (Chi, 2006). Kognitive Belastung kann als ein limitierender Faktor bei der Verarbeitung von visuellem Material aufgefasst werden. Stellt sich in Experten-Novizen-Vergleichen heraus, dass die Experten weniger kognitive Belastung bei der Bearbeitung aufweisen als die Novizen, so kann gefolgert werden, dass die hohe Leistung der Experten darauf zurückzuführen sein kann, dass sie weniger Ressourcen zur Verarbeitung der Aufgabe benötigen. Eine objektive Methode, die kognitive Belastung zu messen, sind Doppeltätigkeitsaufgaben. Studien mit dem so genannten Doppeltätigkeitsparadigma (z.B. Bartmann et al., 1994; Brünken et al., 2003) untersuchen Veränderungen in der Leistung, wenn zusätzlich zu einer Hauptaufgabe (z.B. Fahren) die parallele Bearbeitung einer Nebenaufgabe (kognitiv und/ oder motorisch) verlangt wird. Dieses Forschungsparadigma wurde bereits zur Untersuchung der kognitiven Belastung von Experten und Novizen bei Reaktionszeitaufgaben im Verkehrskontext eingesetzt. In einer Untersuchung von Sagberg (2006) wurde zusätzlich zum Einfluss von Fahrerfahrung auf die Leistung im *Hazard-Perception-Test* untersucht, wie sich eine erhöhte mentale Belastung auf die Gefahrenwahrnehmung auswirkt. Dabei wurde angenommen, dass Fahrnovizen durch eine

Nebenaufgabe stärker in ihrer Leistung im Gefahrenwahrnehmungstest beeinträchtigt würden als Fahrexperten. In einer Längsschnittstudie mit drei Messzeitpunkten wurde den Teilnehmern Videomaterial mit mehreren potenziellen Gefahrensituationen dargeboten, auf die die Probanden innerhalb eines definierten Zeitfensters mit Tastendruck reagieren sollten. Während die Personen einen Aufgabenblock bearbeiteten, sollte die Hälfte der Probanden als Nebentätigkeit einstellige Zahlen addieren. Es zeigten sich generell eher geringe Zusammenhänge zwischen Fahrerfahrung und Leistung im Gefahrenwahrnehmungstest. Außerdem konnten keine signifikanten Unterschiede in der Leistung im *Hazard-Perception-Test* zwischen der Bedingung mit und ohne Nebentätigkeit festgestellt werden. Die durchschnittliche Zahl der Fehler bei der Nebentätigkeit unterschied sich ebenfalls nicht bedeutsam zwischen den Experten und den Novizen. Tendenziell ließ sich aber feststellen, dass die Experten durch die Nebenaufgabe weniger bei der Bearbeitung des Gefahrentests beeinträchtigt wurden und damit geringer mental belastet zu sein schienen, als die Novizen. In dieser Studie zeigte sich demnach kein deutlicher Effekt des Faktors Expertise. Dieser Befund kann auch möglicherweise den schwachen Interaktionseffekt begründen. Kritisch anzumerken ist, dass es sich beim Fahren um einen komplexen Prozess handelt, dem die eingesetzten Reaktionsaufgaben eventuell nicht gerecht wurden. So konnten Bartmann et al. (1994) demgegenüber zeigen, dass die Leistung in einer Nebenaufgabe – in diesem Fall dem taktmäßigen Sprechen von bedeutungslosen Silben – sinkt, sobald die Hauptaufgabe – das Fahren – anspruchsvoller wurde. Fahranfänger waren dabei etwas stärker beansprucht als erfahrene Fahrer. Die Befundlage zur Reaktionszeitmessung mit Nebenaufgaben ist demnach nicht eindeutig.

Soll experimentell überprüft werden, ob die geringere Leistung der Novizen in den Reaktionszeitaufgaben darin begründet liegt, dass sie kognitiv stärker belastet sind als die Experten, besteht auch eine Schwierigkeit darin, geeignete Nebenaufgaben zu identifizieren. Während in den vorgestellten Studien Nebenaufgaben gestellt wurden, die in Bezug auf die Hauptaufgabe als unspezifisch betrachtet werden können, ist aus Gründen der ökologischen Validität der Studien zu erwägen, fahrspezifische Nebenaufgaben zu wählen. Diese müssten dazu allerdings in eine standardisierte Form gebracht werden. Hier besteht noch erheblicher Forschungsbedarf.

Neben subjektiven Methoden, die kognitive Belastung bei der Bearbeitung von Multimediamaterial messen, wie z.B. der am häufigsten eingesetzten Ratingskala von Paas (1992), gewinnen weitere objektive Messmethoden zunehmend an Bedeutung, die die mentale

Belastung mitunter direkt messen können (vgl. z.B. Park, 2010). Vielversprechend erscheint hier der Einsatz von Methoden zur Blickbewegungsmessung, die es ermöglichen zusätzlich die Pupillendilatation zu erfassen. Erste Untersuchungen wurden mit dieser Methode bereits im Verkehrskontext durchgeführt (Schwalm, 2008). Dabei konnte anhand der Erweiterung der Pupille bei verschiedenen computerbasierten verkehrsspezifischen Aufgaben ein Maß berechnet werden – der sogenannte *Index of Cognitive Activity (ICA, Marshall, 2002)* – mit dem zuverlässig solche Verkehrsanforderungen identifiziert werden konnten, die kognitiv belastend waren. Der Einsatz dieser Methode in Experten-Novizen-Vergleichen, würde es ermöglichen Aussagen darüber zu treffen, welche Anforderungen im Verkehrskontext für Novizen besonders beanspruchend sind.

6.1.2 Merkmale des Erwerbs von Fahrexpertise

Zur Untersuchung der Entwicklung von Fahrexpertise wurde gemäß den Empfehlungen des dritten Stadiums des *expert performance approach* (Ericsson & Smith, 1991a) eine Längsschnittuntersuchung durchgeführt. Derzeit fehlt noch ein allgemeines Modell des Expertiseerwerbs für das Autofahren, das zur Systematisierung der Forschung in diesem Bereich dienen könnte. Daher müssen zunächst andere, generelle Modelle der Expertiseentwicklung herangezogen werden, um den Erwerb von Fahrexpertise zu erklären. Anhand der Ergebnisse der vorliegenden Arbeit kann eingeschätzt werden, inwiefern diese Modelle dazu angemessen erscheinen.

In vielen Domänen hat sich die generelle Annahme bestätigt, dass die Entwicklung vom Novizen zum Experten mindestens zehnjähriger intensiver Übung bedarf (Ericsson et al., 1993). In der vorliegenden Arbeit wird der Begriff Expertise nicht im Sinne von Ericsson als außergewöhnliche Leistungsfähigkeit weniger Personen aufgefasst. Um die Erkenntnisse aus der Expertiseforschung für Vorhersagen zur Entwicklung von Alltagsfähigkeiten wie dem Fahren nutzbar zu machen, kann Expertise als eine Dimension betrachtet werden. Beim Erwerb von Expertise steigert eine Person bei ihrer Entwicklung vom Novizen zum Experten ihre Leistung, wobei unterschiedlich viele Stadien benannt werden können (z.B. Dreyfus, S. E. & Dreyfus, 1980; Hoffman, 1998). Abhängig von der zu untersuchenden Domäne hat die Leistungsfähigkeit, die für die verschiedenen festgelegten Stadien charakteristisch ist, eine unterschiedliche praktische Bedeutung. Der Experte im Sinne von Ericssons Expertisebegriff hat für das Fahren im Alltag eine geringe Bedeutung. Das wichtigste Kriterium für Fahrexpertise ist das individuelle Unfallrisiko. Demnach scheint ein sicherheitsrelevantes Maß an Fahrexpertise bereits nach zwei Jahren vorzuliegen, denn mit weiterer Fahrerfahrung

reduziert sich das Risiko zu verunfallen nicht mehr deutlich (Schade, 2001). Nach etwa 5000 Kilometern eigenständigen Fahrens gilt das Unfallrisiko bereits als halbiert (Sagberg, 2002). Diese Befunde stützen die Annahme, dass sich Expertise durch Erfahrung ausbildet. Allerdings hat sich in den Quer- und Längsschnittuntersuchungen dieser Arbeit gezeigt, dass Fahranfänger und erfahrene Fahrer eine ähnlich umfangreiche Wissensbasis aufzuweisen scheinen. Lediglich innerhalb der Fahrausbildung scheinen die Novizen Verkehrswissen zu erwerben. Leistungsunterschiede zwischen den Experten und den Novizen wurden nur dann deutlich, wenn in den Reaktionszeitaufgaben die situationsgerechte Anwendung des Wissens auf simulierte Verkehrssituationen erforderlich wurde. Den Novizen gelang es zudem innerhalb ihrer Fahrausbildung nicht, sich in den Reaktionszeitaufgaben bedeutsam zu verbessern. Diese Ergebnisse weisen darauf hin, dass verschiedene Arten von Verkehrswissen mit unterschiedlicher Geschwindigkeit und unter bestimmten Bedingungen erlernt werden. Innerhalb des Fahrschulunterrichts können deklarative Wissensinhalte erworben werden. Der Transfer des gelernten Wissens und damit die schnelle situationsgerechte Reaktion auf Hinweisreize wird anscheinend erst später beim selbstständigen Fahren erlernt. Diese Befunde lassen sich auf der Basis des ACT*-Modells erklären (Anderson, 1982). Wissen liegt dabei zu Beginn des Fertigkeitserwerbes zunächst in einer deklarativen Form als Faktenwissen vor. Die Fahrschüler befinden sich somit in dieser ersten Phase des Fertigkeitserwerbes, was sich in ihrer guten Leistung in den Wissens- und Regelfragen im Multiple-Choice-Format zeigt, die sie spätestens nach der theoretischen Fahrausbildung zeigen. Durch den wiederholten Einsatz des Wissens in der Kompilierungsphase – also durch die Anwendung der Kenntnisse beim tatsächlichen Fahren – wird das Wissen prozeduralisiert und zunehmend auch automatisiert. Die Automatisierung zeigt sich darin, dass die Experten seltener Fehler machen und früher auf Hinweisreize innerhalb der Szenarien reagieren. Dieser Prozess der Prozeduralisierung von Wissen soll insbesondere beim Aufbau von Routinehandlungen relevant sein. Integriert wird dieser Ansatz in Rasmussens Modell der Handlungskontrolle (1986), das auf das Autofahren bezogen werden kann. Das Modell, besagt, dass unterschiedliche Teilaspekte des Fahrens auf unterschiedlichen kognitiven Ebenen kontrolliert werden. Bei denjenigen Aufgaben im Reaktionstest, in denen die Kenntnis von Vorschriften notwendig war, lässt sich leicht nachvollziehen, dass eine routinierte Reaktion der Experten erfolgen kann, da die Einhaltung bestimmter Regeln im Verkehr alltäglich relevant wird. Hingegen erscheint die Reaktion auf Situationen, wie sie beim Reaktionstest in Szenarien mit explizitem Gefahrenreiz auftraten, schlecht automatisierbar, da sie sehr selten beim Fahren

auftreten. Um zu entscheiden, ob die Reaktion auf Hinweise innerhalb der Szenarien bei den Experten automatisiert erfolgt, bedarf es weiterer Studien, die z.B. mithilfe von Nebentätigkeiten erfassen, in welchem Ausmaß kognitive Ressourcen beansprucht werden. Außerdem sollte überprüft werden, inwiefern die Entscheidungsfindung von den Experten verbalisiert werden kann. Daraus könnten Schlüsse auf eine (un)bewusste Verarbeitung gezogen werden. Noch differenzierter müsste der Ansatz der fünf Stadien von Dreyfus und Dreyfus (1980) überprüft werden. Dabei geht pro Stadium eine von vier kognitiven Funktionen von ihrer primitiven in eine weiterentwickelte ressourcensparende Form über. Lassen sich diese Modelle nicht empirisch bestätigen, sollten alternative Erklärungsansätze, die anstatt einer Simplifizierung der kognitiven Prozesse eher eine Ausdifferenzierung vorhersagen, geprüft werden. In Schema- und Skriptansatz (Minsky, 1975 bzw. Schank, 1982) wird erwartet, dass sich mit der Erfahrung ein differenziertes Wissensnetz ausbildet, in dem auch situationsadäquate Verhaltensimplikationen abgespeichert sind, was das Problemlösen erleichtert und beschleunigt.

Zusammenfassend lässt sich sagen, dass die Studien der vorliegenden Arbeit gezeigt haben, dass Fahrschüler während ihrer Fahrausbildung zwar deklaratives Faktenwissen ausbilden, die schnelle sichere Anwendung dieses Wissens aber innerhalb dieser Zeit nicht erlernen. Sie erreichen dabei nicht nur das Expertenniveau nicht, ihre Leistungen verbessern sich nicht einmal deutlich. Dies bedeutet, dass zur Steigerung dieser Fähigkeiten unbedingt selbstständiges Fahren notwendig ist. Nach Spiro et al. (1988) ist dazu nicht nur, wie in Rasmussens Modell (1986) vorgeschlagen, ein umfangreicher Fahrerfahrungsaufbau (*Quantität*) notwendig sondern auch ein besonders reichhaltiger (*Qualität*), bei dem die Fahranfänger ihr Wissen in diversen Situationen anwenden. Um diese Annahme zu überprüfen, wäre es notwendig, Fahranfänger in ihren ersten beiden Jahren des selbstständigen Fahrens längsschnittlich zu untersuchen und Quantität und Qualität ihrer Erfahrungen als Fahrer im Straßenverkehr möglichst experimentell zu variieren. Durch diese Vorgehensweise könnten kritische Faktoren bei der Ausbildung von Gefahrenwahrnehmung und allgemeinem Verkehrsüberblick identifiziert werden.

Weiterführend bleibt festzuhalten, dass es sich beim Erwerb von Fahrexpertise um einen Prozess handelt, bei dem verschiedene Teilfertigkeiten erworben werden, die Kognitionen, Affekte und auch die Psychomotorik einschließen (Leutner et al., 2009). Die in dieser Studie zur Messung von Fahrexpertise eingesetzten Aufgaben erfassen jedoch lediglich kognitive Aspekte des Fahrens. Beispielsweise konnte in Vorstudie 2 kein Zusammenhang

zwischen der Leistung in den Reaktionszeitaufgaben und der Tendenz zu Sensationssuche nachgewiesen werden. Die Entwicklung affektiver und psychomotorischer Fähigkeiten, die sich eventuell unabhängig von kognitiven Fähigkeiten ausbilden, kann mit diesen Aufgaben nicht verfolgt werden. Annahmen über ihre Entwicklung sollten dennoch in einem Entwicklungsmodell der Fahrexpertise getroffen werden. Eine Überprüfung der Gültigkeit eines solchen Modells erfordert wiederum Instrumente, die diese Aspekte der Fahrexpertise angemessen erfassen können.

6.2 Bedeutung der Ergebnisse für die Praxis der Fahrerlaubnisprüfung und Fahrausbildung

Neben dem Ziel, Aussagen zu Merkmalen der Fahrexpertise und deren Entwicklung zu treffen, wurde mit dieser Arbeit die Absicht verfolgt, empirisch gestützte Prinzipien für die Gestaltung der Fahrerlaubnisprüfung zu erarbeiten. Aus den Erkenntnissen ergeben sich ebenfalls Implikationen für die Beschaffenheit der Fahrausbildung.

6.2.1 Bedeutung für die Gestaltung der Fahrerlaubnisprüfung

Mithilfe eines kriteriumsorientierten Tests wie z.B. der Fahrerlaubnisprüfung werden Können von Nicht-Könnern getrennt (Klauer, K. J., 1987). Aufgaben, die dies ermöglichen sind als kriteriumsvalid zu bezeichnen und ihr Einsatz in der Prüfung kann als gerechtfertigt betrachtet werden. Innerhalb der vorliegenden Arbeit wurden in Vorstudie 1 sowie in der Längsschnittstudie Aufgaben eingesetzt, wie sie in ähnlicher Form in der derzeitigen theoretischen Fahrerlaubnisprüfung vorkommen: die Wissens- und Regelaufgaben im Multiple-Choice-Antwortformat. Die Ergebnisse der Längsschnittstudie widerlegen die Kriteriumsvalidität dieser Aufgaben, da die Leistung der Novizen nach Abschluss ihrer Fahrausbildung der Leistung der Experten in diesen Aufgaben entsprach. Reelle Leistungsunterschiede zwischen erfahrenen Fahrern und unerfahrenen Fahrern sollten zu diesem Zeitpunkt allerdings noch bestehen, was sich im gesteigerten initialen Unfallrisiko der Fahranfänger ausdrückt (vgl. Schade, 2002). Möglicherweise handelt es sich bei deklarativem Verkehrswissen um eine notwendige aber keine hinreichende Voraussetzung für das sichere Fahren. Dies würde bedeuten, dass es zwar in der Fahrerlaubnisprüfung abgeprüft werden sollte, Experten-Novizen-Unterschiede aber nicht zwingend vorliegen müssten, um von einer validen Testung ausgehen zu können.

Wie bereits in Kapitel 6.1 dargestellt wurde, könnten der fehlende Einfluss des Faktors Expertise auch weniger mit dem Inhalt der Aufgaben als mit der Form der Überprüfung

zusammenhängen. Daher könnte ein Grund dafür, dass die Aufgaben Fahrpraxis nicht abzubilden vermögen, die mangelnde Handlungsnähe des Multiple-Choice-Antwortformates sein. Während die bisherigen Aufgaben zur Überprüfung von Fahrexpertise nicht geeignet erscheinen, sollte jedoch auf das Abprüfen von verkehrsrelevantem Wissen nicht verzichtet werden. Daher liegt es nahe, weiterführend neben dynamischen Präsentationsformaten auch den Einsatz innovativer Antwortformate mit größerer Handlungsnähe zu prüfen. Die Ergebnisse der Vorstudie 2 und der Längsschnittstudie belegen, dass der Reaktionstest eine valide Messung eines Aspektes von Fahrkompetenz ermöglicht. Daher kann eine Erfassung von Verkehrswissen und Gefahrenwahrnehmung mithilfe von Aufgaben, die den entwickelten Reaktionszeitaufgaben entsprechen, prinzipiell empfohlen werden.

Wenn für die Kriteriumsvalidität der Aufgaben tatsächlich die Handlungsnähe, bzw. ökologische Validität des Antwortformates relevant ist, könnte man mithilfe einer weiteren Steigerung der Handlungsnähe eventuell noch validere Entscheidungen treffen. Noch näher an der Realität des Fahrens sind Simulatorfahrten, in denen sich die Testteilnehmer selbstständig als Fahrer durch eine virtuelle Welt bewegen. Dieses Format wird bisher nicht in der Fahrerlaubnisprüfung eingesetzt. Aktuelle Fahrsimulatoren dienen eher dem Training fahrrelevanter Fertigkeiten, werden aufgrund hoher Kosten aber kaum innerhalb der regulären Fahrausbildung in Deutschland genutzt (Weiß, Petzold, Bannert & Krems, 2007). Underwood, Geoffrey, Crundall und Chapman (2011) erachten Fahrsimulatoren als vielversprechend zur Leistungsmessung. Die Autoren führen allerdings an, dass Fahrsimulatoren bisher erst in Bezug auf basale perzeptuell-motorische Fähigkeiten (z.B. Geschwindigkeitskontrolle, Spurhalten) validiert wurden. Eine Validierung bezüglich kognitiv anspruchsvollerer Fähigkeiten (z.B. Gefahrenwahrnehmung) steht bisher noch aus. Bevor Simulatoren zur Leistungsmessung eingesetzt werden können, sollten den Aufgaben der Reaktionstests entsprechend, standardisierte Szenarien und Fahraufgaben entwickelt und validiert werden.

Aufgrund der Vorteile dynamischer Präsentationen bei der Darbietung von Lernmaterial (vgl. Höffler & Leutner, 2007), sollte in den vorgestellten Vorstudien ermittelt werden, welchen Nutzen Animationen gegenüber statischen Darstellungen im Testmaterial haben. Für den Einsatz dynamischer Elemente gibt es verschiedene Möglichkeiten. Momentan wird in Erprobungsuntersuchungen der TÜV | DEKRA arge tp 21 der Einsatz einer Ergänzung der bestehenden Aufgaben der theoretischen Fahrerlaubnisprüfung mit Mehrfachwahlantwortformat durch eine kurze Animation des Verkehrsszenarios geprüft. In

Vorstudie 1 der vorliegenden Arbeit zeigte sich, dass solche Animationen, die einen erklärenden Charakter haben (Wissensaufgaben), die Novizen bei der Bearbeitung der Aufgabe unterstützen. Da ein gesteigertes Instruktionsverständnis Reliabilität und Validität eines Tests positiv beeinflusst, kann empfohlen werden, die Situationen, auf die sich die Testaufgaben beziehen, durch Animationen zu veranschaulichen. In Aufgaben, in denen die bildlichen Darstellungen nicht als Veranschaulichung der Verkehrsgegebenheiten dienen sondern selbst einen wesentlichen Anteil an den Anforderungen der Aufgabe darstellen (Reaktionstest), zeigte sich kein Unterstützungseffekt der Dynamik für die Novizen. Dynamische Darstellungen erschwerten die Bearbeitung der Aufgaben für die Novizen und die Experten in gleichem Maße. Die Tatsache, dass die Anforderungen in der dynamischen Version des Testmaterials eher den Anforderungen beim realen Fahren entsprechen – es muss in einem kritischen Zeitfenster reagiert werden – sprechen für die ökologische Validität der Animationen und prinzipiell für deren Einsatz in der Fahrerlaubnisprüfung. Allerdings gilt bei Letzterem zu bedenken, dass sich in der Längsschnittuntersuchung gezeigt hat, dass die Fähigkeiten, die zur Lösung der Reaktionszeitaufgaben notwendig sind, nicht während der Fahrausbildung entwickelt werden. Demnach wäre das Abprüfen dieser Fähigkeiten zu einem Zeitpunkt, an dem der Bewerber noch über kaum Fahrerfahrung verfügt, unangebracht. Sinnvoller wäre es, den Reaktionstest nach einer Phase des eigenständigen Fahrens durchzuführen. Zur Identifikation eines geeigneten Zeitpunktes dafür wäre ein empirisch gesichertes Modell des Fahrkompetenzerwerbs wünschenswert. Dies ermöglicht sogenannte Stufenführerscheine, die eine Passung der jeweiligen Fahrkompetenzstufe zu den Rechten der Führerscheinneulinge, die ihnen erst nach und nach zugestanden werden, gewährleisten (Willmes-Lenz, 2010). Erstrebenswerter wäre es allerdings, die Fahrausbildung so anzupassen, dass alle relevanten Fähigkeiten bereits dabei erworben werden und das Anfängerrisiko somit minimiert würde.

6.2.2 Bedeutung für die Gestaltung der Fahrausbildung

Ändert man die Anforderungen einer Prüfung, hat dies in der Regel Auswirkungen auf die Gestaltung der Ausbildungsphase, in der nicht zuletzt die optimale Vorbereitung auf die Prüfung verfolgt wird. Auf Grund des hohen Unfallrisikos, dem Fahranfänger in ihren ersten Monaten des selbstständigen Fahrens ausgesetzt sind, kann man vermuten, dass sie durch die Fahrausbildung nicht optimal vorbereitet werden. Es hat sich in den Studien dieser Arbeit wie auch in früheren Studien gezeigt, dass besonders bei anspruchsvollen kognitiven Aspekten des Fahrens, wie z.B. bei der Gefahrenwahrnehmung, der Aufmerksamkeit und visuellen

Strategien beim Absuchen der Umwelt, Leistungsunterschiede zwischen Experten und Novizen bestehen. Geht man nun weiterhin davon aus, dass diese kognitiven Aspekte ausschlaggebend für das Unfallrisiko sind, lässt sich daraus ableiten, dass in einer optimierten Fahrausbildung genau diese Aspekte stärker fokussiert und trainiert werden sollten. Expertise wird in der Mehrzahl der Domänen als von den meisten Individuen erwerbbar betrachtet (Ericsson & Smith, 1991a). Studien, in denen Expertise relativ als Leistungsunterschied zwischen Experten und Novizen definiert wird, liefern demnach nicht nur Erklärungen und Beschreibungen von Expertenwissen und -können, sondern zeigen auf, wie Novizen zu Experten werden können (Chi, 2006). Demnach kann man sich bei der Suche nach geeigneten Trainingsmaßnahmen auf generelle Modelle des Expertiseerwerbs beziehen. Es hat sich gezeigt, dass der Expertiseerwerb besonders durch *deliberate practice*, also durch planvolle eigenständige Übung, begünstigt wird (Ericsson et al., 1993). Diese besondere Art der Übung findet aber in der derzeitigen Fahrausbildung kaum statt. Allein für die Aufgaben der theoretischen Fahrerlaubnisprüfung, deren Kriteriumsvalidität in dieser Arbeit nicht bestätigt werden konnte, trainieren die Fahrschüler auf diese Weise. Dadurch erreichen sie am Ende ihrer Fahrausbildung den Kenntnisstand der Experten. *Deliberate practice* wäre aber wahrscheinlich auch für die anderen kognitiven Aspekte des Fahrens, die sich als bedeutsam für Fahrexpertise erwiesen haben, hilfreich. Das Begleitete Fahren ab 17 stellt dabei bereits einen Schritt in die richtige Richtung dar, da die Fahranfänger innerhalb dieser Phase dem Realverkehr durch die Anwesenheit der Begleitperson unter besonders sicheren Bedingungen ausgesetzt sind. Die Wirksamkeit dieser Maßnahme konnte bereits empirisch nachgewiesen werden (Stiensmeier-Pelster, 2007). Gezielt besondere Verkehrsanforderungen zu üben, wie es im *deliberate-practice*-Ansatz gefordert wird, ist allerdings im Realverkehr schwierig zu realisieren. Dazu müssten relevante Verkehrskonstellationen wiederholt aufgesucht werden. Problematisch erscheint es dabei, besonders witterungs- oder straßenbaubedingte Anforderungen herzustellen. Einfacher könnten spezifische kognitive Fähigkeiten durch Computerprogramme oder Fahrsimulatoren trainiert werden, die es erlauben, jegliche Art von Verkehrsgegebenheiten zu simulieren. Mit solchen Methoden könnten mithilfe von *deliberate practice* beispielsweise *Hazard Perception*, Strategieauswahl und die adäquate Verteilung kognitiver Ressourcen trainiert werden (Durso & Dattel, 2006). Innerhalb der vorliegenden Arbeit wurden zur Leistungsmessung bereits Szenarien entwickelt, die den Unterschied zwischen den Experten und den Novizen reproduzieren konnten, auch wenn letztere ihre Fahrausbildung bereits abgeschlossen hatten. Dies zeigte, dass durch die derzeitige

Fahrschulausbildung die für die Lösung dieser Aufgaben erforderlichen Kompetenzen nicht in ausreichendem Maße vermittelt werden. In einer weiterführenden Trainingsstudie könnte empirisch untersucht werden, inwieweit mithilfe eines Trainingsprogramms, das diese Szenarien beinhaltet, die Verkehrswahrnehmung der Fahranfänger im Realverkehr verbessert werden kann.

Innerhalb der vorliegenden Arbeit konnte gezeigt werden, dass gute und weniger gute Methoden vorliegen, um relevante kognitive Aspekte der Fahrexpertise und des Erwerbs von Fahrexpertise zu erfassen. Es bleibt die Forderung nach einem empirisch geprüften, vollständigen Modell der Entwicklung von Fahrexpertise, das die Forschung in diesem anwendungsorientierten Bereich strukturiert, indem es die Evaluation von Methoden zur Erfassung von Fahrexpertise und von Trainingsmaßnahmen ermöglicht. Auf der Grundlage dieser Forschung ist eine enge Verzahnung von Empirie und Praxis wünschenswert, um Instruktion und Assessment stets nach den aktuellen Forschungsimplicationen zu gestalten und letztendlich das bedeutende Ziel einer Verringerung des Unfallrisikos der Fahranfänger zu erreichen.

Literaturverzeichnis

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369-406.
- Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological Review*, 94, 192-210.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Atkinson, R. C. & Shiffrin, R. M. (1969). Storage and retrieval processes in short-term memory. *Psychological Review*, 76, 179-193.
- Bailly, B., Bellet, T. & Goupil, C. (2003). *Drivers' mental representations: experimental study and training perspectives*. Paper presented at the First International Conference on Driving Behaviour and Training, Stratford-upon-avon, UK.
- Ballesteros, M. F. & Dischinger, P. C. (2002). Characteristics of traffic crashes in Maryland (1996-1998): differences among the youngest drivers. *Accident Analysis & Prevention*, 34(3), 279-284.
- Bartmann, A. (1995). *Zur Erfassung von kognitiver Beanspruchung beim Führen von Kraftfahrzeugen. Eine Feldstudie*. Aachen: Skaker.
- Bartmann, A., Debus, G. & Heller, D. (1994). *Routinehandlungen beim Führen von Kraftfahrzeugen* (Forschungsbericht für die Bundesanstalt für Straßenwesen). Aachen: Institut für Psychologie der RWTH Aachen.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W. et al. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Bédard, J. & Chi, M. T. H. (1992). Expertise. *Current Directions in Psychological Science (Wiley-Blackwell)*, 1(4), 135-139.
- Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 287-296). New York: Cambridge University Press.
- Betrancourt, M. & Tversky, B. (2000). Effect of computer animation on users performance: A review. *Le travail Humain*, 63, 311-330.
- Biermann, A. (2007). *Gefahrenwahrnehmung und Expertise - Möglichkeiten der Erfassung und Eignung als Prädiktor des Verunfallungsrisikos junger Fahranfänger: Dissertation, Digitale Bibliothek Thüringen, <http://www.db-thueringen.de/servlets/DocumentServlet?id=9049>*. .
- Biermann, A., Skottke, E.-M., Anders, S., Brünken, R., Debus, G. & Leutner, D. (2008). Entwicklung und Überprüfung eines Wirkungsmodells: Eine Quer- und Längsschnittstudie. In R. Brünken, G. Debus & D. Leutner (Hrsg.), *Wirkungsanalyse und Bewertung der neuen Regelungen im Rahmen der Fahrerlaubnis auf Probe (Berichte der BAST, Heft M 194)* (S. 46-111). Bremerhaven: Wirtschaftsverlag NW.
- Borowsky, A., Shinar, D. & Oron-Gilad, T. (2007). *Age, skill, and hazard perception in driving*. Paper presented at the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Stevenson, Washington.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5). Berlin: Springer-Verlag.

-
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer Medizin Verlag.
- Bromme, R. (1992). *Der Lehrer als Experte*. Bern, Göttingen: Huber.
- Brünken, R., Plass, J. L. & Leutner, D. (2003). Direct Measurement of Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38, 53-61.
- Brünken, R., Steinbacher, S. & Leutner, D. (2000). Räumliches Vorstellungsvermögen und Lernen mit Multimedia. In D. Leutner & R. Brünken (Hrsg.), *Neue Medien in Unterricht, Aus- und Weiterbildung*. Münster: Waxmann.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Cellier, J. M., Eyrolle, H. & Marine, C. (1997). Expertise in dynamic environments. *Ergonomics*, 40(1), 28-50.
- Chan, E., Pradhan, A. K., Pollatsek, A., Knodler, M. A. & Fisher, D. L. (2010). Are driving simulators effective tools for evaluating novice drivers' hazard anticipation, speed management, and attention maintenance skills? *Transportation Research Part F*(13), 343-353.
- Chapman, P. R. & Underwood, G. (1998). Visual search of dynamic scenes: Event types and the role of experience in viewing driving situations. In G. Underwood (Hrsg.), *Eye guidance in reading and scene perception* (S. 369-393). Oxford: Elsevier.
- Charness, N. (1989). Expertise in chess and bridge. In D. Klahr & K. Kotovsky (Hrsg.), *Complex information processing. The impact of Herbert A. Simon* (S. 183-208). Hillsdale, NJ: Erlbaum.
- Charness, N., Tuffiash, M., Krampe, R., Reingold, E. & Vasyukova, E. (2005). The role of deliberate practice in chess. *Applied Cognitive Psychology*, 19, 151-165.
- Chase, W. G. & Ericsson, K. A. (1982). Skilled Memory. In J. R. Anderson (Hrsg.), *Cognitive skills and their acquisition* (Bd. 16, S. 1-58). New York: Academic Press.
- Chase, W. G. & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Hrsg.), *The cambridge handbook of expertise and expert performance* (S. 21-30). New York: Cambridge University Press.
- Chi, M. T. H., Feltovich, P. J. & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science: a Multidisciplinary Journal*, 5(2), pp. 121 - 152.
- Chi, M. T. H., Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Hrsg.), *Advances in the psychology of human intelligence* (Bd. 1). Hillsdale, NJ: Erlbaum.
- Cooper, P. J., Pinili, M. & Chen, W. (1995). An examination of the crash involvement rates of novice drivers aged 16 to 55. *Accident Analysis and Prevention*, 27(1), 89-104.
- Crundall, D. E., Underwood, G. & Chapman, P. R. (1998). How much do novice drivers see? The effects of demand on visual search strategies in novice and experienced drivers. In

-
- G. Underwood (Hrsg.), *Eye guidance in reading and scene perception* (S. 395-417). Oxford: Elsevier.
- Csikszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
- Currie, L. (1969). The perception of danger in a simulated driving task. *Ergonomics*, 12, 841-849.
- Darby, P., Murray, W. & Raeside, R. (2009). Applying online fleet driver assessment to help identify, target and reduce occupational road safety risks. *Safety Science*, 47(3), 436-442.
- de Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton.
- Debus, G., Leutner, D., Brünken, R., Skottke, E.-M. & Biermann, A. (2008). *Wirkungsanalyse und Bewertung der neuen Regelungen im Rahmen der Fahrerlaubnis auf Probe* (Berichte der BASt, Heft M 194). Bremerhaven: Wirtschaftsverlag NW.
- Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research*, 30(4), 225-236.
- Dreyfus, H. L. & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press.
- Dreyfus, S. E. & Dreyfus, H. L. (1980). A five-stage model of the mental activities involved in directed skill acquisition. *Distribution*, (February), 1-18.
- Durso, F. T. & Dattel, A. R. (2006). Expertise and Transportation. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Hrsg.), *The cambridge handbook of expertise and expert performance* (S. 355-371). New York: Cambridge University Press.
- Ekstrom, R. B., French, J. B., Harman, H. H. & Dermen, D. (1978). *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton, New Jersey: Educational Testing Service.
- Elander, J., West, R. & French, D. (1993). Behavioural correlates of individual differences in road-traffic crash risk: an examination of methods and findings. *Psychological Bulletin*, 113, 279-294.
- Ellinghaus, D. & Steinbrecher, J. (1990). *Junge Fahrer - Besser oder schlechter als ihr Ruf?* Köln: UNIROYAL.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors* 37(1), 32-64.
- Endsley, M. R. & Garland, D. J. (2000). *Situation Awareness Analysis and Measurement*: Mallory International.
- Ericsson, K. A. & Crutcher, R. J. (1990). The nature of exceptional performance. In P. B. Baltes, D. L. Featherman & R. M. Lerner (Hrsg.), *Life-span development and behavior* (S. 187-217). Hillsdale, NJ: Erlbaum.
- Ericsson, K. A. & Kintsch, W. (1995). Long-Term Working Memory. *Psychological Review*, 102(2), 211-245.
- Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406.

-
- Ericsson, K. A. & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47(1), 273–305.
- Ericsson, K. A. & Smith, J. (1991a). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Hrsg.), *Toward a general theory of expertise: Prospects and limits* (S. 1-38). New York: Cambridge University Press.
- Ericsson, K. A. & Smith, J. (1991b). *Toward a general theory of expertise: prospects and limits*. Cambridge: Cambridge University Press.
- Ericsson, K. A. & Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Hrsg.), *Complex information processing: The impact of Herbert A. Simon* (S. 235-267). Hillsdale, NJ: Erlbaum.
- Ericsson, K. A. & Ward, P. (2007). Capturing the Naturally Occurring Superior Performance of Experts in the Laboratory: Toward a Science of Expert and Exceptional Performance. *Current Directions in Psychological Science (Wiley-Blackwell)*, 16(6), 346-350.
- Evans, L. (1991). *Traffic safety and the driver*. New York: Van Nostrand Reinhold.
- Farrington-Darby, T. & Wilson, J. R. (2006). The nature of expertise: A review. *Applied Ergonomics*, 37(1), 17-32.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Fitts, P. M. & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks Cole.
- Friege, G. (2000). *Wissen und Problemlösen. Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf Grundlage des Experten-Novizenvergleichs*. (Bd. Band 19). Berlin: Logos Verlag.
- Galton, F., Sir. (1969/1979). *Hereditary genius: An inquiry into it's laws and consequences*. (Originally published in 1969). London: Julian Friedmann.
- Garner, R., Alexander, P. A., Gillingham, M. G., Kulikowich, J. M. & Brown, R. (1991). Interest and learning from text. *American Educational Research Journal*, 28, 643-659.
- Garner, R., Gillingham, M. G. & White, C. S. (1989). Effects of “seductive details” on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, 6(1), 41-57.
- Gentner, D. R. (1988). Expertise in typewriting. In M. T. H. Chi, R. Glaser & M. J. Farr (Hrsg.), *The nature of expertise* (S. 1-21). Hillsdale, NJ: Erlbaum.
- Glaser, R. & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser & M. J. Farr (Hrsg.), *The nature of expertise* (S. xv-xxviii). Hillsdale: NJ: Erlbaum.
- Gniech, G., Oetting, T. & Brohl, M. (1993). Untersuchungen zur Messung von "Sensation-seeking". *Bremer Beiträge zur Psychologie, Nr. 110*, 3-60.
- Gobet, F. (1996). Expertise und Gedächtnis. In H. Gruber & A. Ziegler (Hrsg.), *Expertiseforschung. Theoretische und methodische Grundlagen* (S. 58-79). Opladen: Westdeutscher Verlag.

-
- Grattenthaler, H. & Krüger, H.-P. (2009). *Bedeutung der Fahrpraxis für den Kompetenzerwerb beim Fahrenlernen*. Bergisch Gladbach: Bundesanstalt für Straßenwesen.
- Grayson, G. B., Maycock, G., Groeger, J. A., Hammond, S. M. & Field, D. T. (2003). *Risk, hazard perception and perceived control* (TRL Report 560). Crowthorne, Berkshire: TRL Limited.
- Grayson, G. B. & Sexton, B. F. (2002). *The development of hazard perception testing* (TRL Report 558). Crowthorne, Berkshire: TRL Limited.
- Gregersen, N. P. (1996). Young drivers' overestimation of their own skill - an experiment on the relation between training strategy and skill. *Accident Analysis and Prevention*, 28(2), 243-250.
- Gregersen, N. P., Berg, H. Y., Engström, I., Nolén, S., Nyberg, A. & Rimmö, P. A. (2000). Sixteen years age limit for learner drivers in Sweden - an evaluation of safety effects. *Accident Analysis and Prevention*, 32(1), 25-35.
- Groeger, J. A. (2000). *Understanding driving. Applying cognitive psychology to a complex everyday task*. Hove, UK: Psychology Press.
- Gruber, H. (1994). *Expertise. Modelle und empirische Untersuchungen*. Opladen: Westdeutscher Verlag.
- Gruber, H. & Lehman, A. C. (2007). *Entwicklung von Expertise und Hochleistung in Musik und Sport (Forschungsbericht Nr. 26)*. Regensburg: Universität Regensburg, Lehrstuhl für Lehr-Lernforschung.
- Gruber, H. & Mandl, H. (1995). *Auswirkungen von Erfahrungen auf die Entwicklung von Expertise (Forschungsbericht Nr.45)*. München: Ludwig-Maximilians-Universität.
- Gruber, H. & Mandl, H. (1996). Das Entstehen von Expertise. In J. Hoffmann & W. Kintsch (Hrsg.), *Enzyklopädie der Psychologie, Serie Kognition, Band Lernen* (S. 583-615). Göttingen: Hogrefe.
- Gruber, H., Renkl, A. & Schneider, W. (1994). Expertise und Gedächtnisentwicklung. Längsschnittliche Befunde aus der Domäne Schach. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 26, 53-70.
- Hall, J. & West, R. (1996). The role of formal instruction and informal practice in learning to drive. *Ergonomics*, 39, 693-706.
- Harp, S. F. & Mayer, R. E. (1998). How Seductive Details Do Their Damage: A Theory of Cognitive Interest in Science Learning. *Journal of Educational Psychology*, 90(3), 414-434.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127-143). Heidelberg: Springer.
- Hatakka, M., Keskinen, E., Baughan, C., Goldenberg, C., Gregersen, N. P., Groot, H. et al. (2003). *BASIC driver training: New models. EU-project Final Report*: University of Turku, Finland.
- Hatakka, M., Keskinen, E., Gregersen, N. P. & Glad, A. (1999). Theories and aims of educational and training measures. In S. Siegrist (Hrsg.), *Driver training, testing and licensing - towards theory-based management of young drivers' injury risk in road traffic. Results of EU-project GADGET* (S. 13-48). Bern: BFU.

-
- Hegarty, M. & Kriz, S. (2008). Effects on knowledge and spatial ability on learning from animation. In R. Lowe & W. Schotz (Hrsg.), *Learning with animation. Research implications for design* (S. 3-29). New York: Cambridge University Press.
- Hodges, N. J., Huys, R. & Starkes, J. L. (2007). Methodological review and evaluation of research in expert performance in sport. In G. Tenenbaum & R. C. Eklund (Hrsg.), *Handbook of Sport Psychology* (Bd. 3, S. 161-183). Hoboken, NJ: John Wiley & Sons.
- Höffler, T. N. (2010). Spatial Ability: Its influence on learning with visualizations - a meta-analytic review. *Educational Psychology Review*, 22, 245-269.
- Höffler, T. N. (2011). The role of spatial ability in learning from instructional animations - Evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior*, 27, 209-216.
- Höffler, T. N. & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17, 722-738.
- Hoffman, R. R. (1998). How can expertise be defined?: Implications of research from cognitive psychology In R. Williams, W. Faulkner & J. Fleck (Hrsg.), *Exploring Expertise* (S. 81-100). New York: Macmillan.
- Horswill, M. S. & McKenna, F. P. (1999). The development, validation, and application of a video-based technique for measuring an every day risk-taking behaviour: Drivers' speed choice. *Journal of Applied Psychology*, 84(6), 977-985.
- Horswill, M. S. & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Hrsg.), *A cognitive approach to situation awareness: Theory and application* (S. 155-175). Aldershot: Ashgate.
- Hosking, S. G., Liu, C. C. & Bayly, M. (2010). The visual search patterns and hazard responses of experienced and inexperienced motorcycle riders. *Accident Analysis and Prevention*, 42, 196-202.
- Huestegge, L., Skottke, E.-M., Anders, S., Muesseler, J. & Debus, G. (2010). The development of hazard perception: Dissociation of visual orientation and hazard perception. *Transportation Research*, 13F, 1-8.
- Hull, M. A. & Christie, R. J. (1993). *The hazard perception test: The Geelong trial and future developments* (VicRoads Report GR 93-13).
- Ivers, R. Q., Blows, S. J., Stevenson, M. R., Norton, R. N., Williamson, A., Eisenbruch, M. et al. (2006). A cohort study of 20822 young drivers: the DRIVE study methods and population. *Injury Prevention*, 12, 385-389.
- Jacobson, M. J. & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, 12, 301-333.
- Jarodzka, H., Scheiter, K., Gerjets, P. & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, 20(2), 146-154.
- Jonah, B. A. (1986). Youth and traffic accident risk: Possible causes and potential solutions. *Accident analysis and prevention*, 18(4), 253-254.

-
- Jonah, B. A. (1996). Sensation seeking and risky driving. In T. Rothengatter & E. C. Vaya (Hrsg.), *Traffic and transport psychology. Theory and Application* (S. 259-267). Amsterdam: Pergamon.
- Kahneman, D. & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman, D. R. Davies & J. Beatty (Hrsg.), *Variants of attention* (S. 29-61). New York: Academic Press.
- Kalyuga, S. (2005). Prior knowledge principle in multimedia learning. In R. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 325-337).
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23-32.
- Katayama, T., Motoki, M., Ochiai, H. & Nakanishi, M. (1991). Comparison of riding behaviour between inexperienced riders and experienced riders. *Forschungshefte Zweiradsicherheit*, 7, 383-403.
- Keskinen, E. (1996). Warum ist die Unfallrate junger Fahrer und Fahrerinnen höher? In BAST (Hrsg.), *Junge Fahrer und Fahrerinnen* (Bd. Berichte der BAST, Heft M 52, S. 42-55). Bremerhaven: Wirtschaftsverlag NW.
- Klauer, K. J. (1987). *Kriteriumsorientierte Tests*. Göttingen: Hogrefe.
- Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D. & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/ crash risk: An analysis using the 100-Car Naturalistic Driving Study data* (No. DOT HS 810 594). Washington, DC: National Highway Traffic Safety Administration.
- Klieme, E., Maag-Merki, K. & Hartig, J. (2007). Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (S. 5-16). Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Kolodner, J. L. (1983). Towards an understanding of the role of experience in the evolution from novice to expert. *International Journal of Man-Machine Studies* 19, 497-518.
- Kraftfahrt-Bundesamt. (2010). Verkehrsauffälligkeiten - Die „Punkte in Flensburg“.
- Krampe, R. & Charness, N. (2006). Aging and expertise. In K. A. Ericsson, N. Charness, P. Feltovich & R. R. Hoffmann (Hrsg.), *Cambridge handbook of expertise and expert performance* (S. 723-742). Cambridge, UK: Cambridge University Press.
- Krems, J. (1994). *Wissensbasierte Urteilsbildung: diagnostisches Problemlösen durch Experten und Expertensysteme*. Bern: Huber.
- Kroj, G. & Schulze, H. (2002). Das Unfallrisiko junger Fahrerinnen und Fahrer - Ursachen und Lösungsperspektiven. In BAST (Hrsg.), *Zweite Internationale Konferenz "Junge Fahrer und Fahrerinnen"* (Bd. Berichte der BAST, Heft M 143, S. 21-28). Bremerhaven: Wirtschaftsverlag NW.
- Laapotti, S., Keskinen, E., Hatakka, M., Hernetkoski, K., Katila, A., Peräaho, M. et al. (2006). Driving circumstances and accidents among novice drivers *Traffic Injury Prevention*, 7, 237-237.
- Laapotti, S., Keskinen, E., Hatakka, M. & Katila, A. (2001). Novice drivers' accidents and violations - a failure on higher or lower hierarchical levels of driving behaviour. *Accident Analysis and Prevention*, 33, 759-769.

-
- Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D. & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. H. Chi & R. Glaser (Hrsg.), *The Nature of Expertise* (S. 311-342). Hillsdale, NJ: Erlbaum.
- Leutner, D. & Brünken, R. (2002). Lehr-lernpsychologische Grundlagen des Erwerbs von Fahr- und Verkehrskompetenzen. In BAST (Hrsg.), *Zweite Internationale Konferenz "Junge Fahrer und Fahrerinnen"* (Bd. Berichte der BAST, Heft M 142, S. 76-87). Bremerhaven: Wirtschaftsverlag NW.
- Leutner, D., Brünken, R. & Willmes-Lenz, G. (2009). Fahren lernen und Fahrausbildung. In H.-P. Krüger (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete, Serie VI Verkehrspsychologie, Bd. 2 Anwendungsfelder der Verkehrspsychologie* (S. 1-79). Göttingen: Hogrefe.
- Lowe, R. & Schnotz, W. (Hrsg.). (2008). *Learning with animation*: Cambridge University Press.
- Malone, S., Biermann, A., Buch, S. & Brünken, R. (2012). *Neue Aufgabenformate in der Theoretischen Fahrerlaubnisprüfung* (Berichte der BAST, Heft M 222). Bremerhaven: Wirtschaftsverlag NW.
- Marshall, S. P. (2002). *The Index of Cognitive Activity: Measuring Cognitive Workload*. Paper presented at the Conference on Human Factors and Power Plants, Scottsdale, Arizona.
- Maycock, G., Lockwood, C. R. & Lester, J. F. (1991). *The accident liability of car drivers* (TRL Report 275). Crowthorne, Berkshire: TRL Limited.
- Mayer, R. E. (2001). *Multimedia Learning*. Cambridge, UK: Cambridge University Press.
- Mayer, R. E. (2008). Research-based principles for learning with animation. In R. Lowe & W. Schnotz (Hrsg.), *Learning with animation. Research implications for design* (S. 30-48). Cambridge: Cambridge University Press.
- Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth ten thousand words? Extensions of a dual-coding theory of multimedia learning *Journal of Educational Psychology*, 86, 389-401.
- Mayhew, D. R. & Simpson, H. M. (1995). *The role of driving experience: Implications for the training and licensing of new drivers*. Toronto, Ontario: Insurance Bureau of Canada.
- Mayhew, D. R. & Simpson, H. M. (1996). *Effectiveness and role of driver education and training in a graduated licensing system*. Ottawa, Canada: Traffic Injury Research Foundation.
- Mayhew, D. R., Simpson, H. M. & Pak, A. (2003). Changes in collision rates among novice drivers during the first months of driving. *Accident Analysis and Prevention*, 35, 683-691.
- McKenna, F. P. & Crick, J. L. (1991). Experience and expertise in hazard perception. In G. B. Grayson & J. F. Lester (Hrsg.), *Behavioural research in road safety*. Crowthorne, Berkshire: Transport and Road Research Laboratory.
- McKenna, F. P. & Crick, J. L. (1994a). *Developments in hazard perception. Final Report*: Department of Transport (UK).
- McKenna, F. P. & Crick, J. L. (1994b). *Hazard perception in drivers: A methodology for testing and training* (Contractor Report 313). Crowthorne, Berkshire: TRL Limited.

-
- McKenna, F. P. & Horswill, M. S. (1999). Hazard perception and its relevance for driver licensing. *IATSS Research*, 23(1), 36-41.
- McKenna, F. P. & Horswill, M. S. (2006). Risk taking from the participant's perspective: The case of driving and accident risk. *Health Psychology*, 25, 163-170.
- McKnight, A. J. & McKnight, S. A. (2003). Young Novice Drivers: Careless or Clueless. *Accident Analysis and Prevention*, 35, 921-925.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Mills, K. L., Hall, R. D., McDonald, M. & Rolls, G. (1998). *The effects of hazard perception training on the development of novice driver skills* (Road Safety Report No. 4). London: Department of the Environment, Transport and the Regions.
- Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Hrsg.), *The psychology of computervision* (S. 211-277). New York: McGraw-Hill.
- Moe, D. (1996). Junge Fahrerinnen und Fahrer, Risikoverhalten und Unfälle. In BAST (Hrsg.), *Junge Fahrer und Fahrerinnen* (Bd. Berichte der BAST, Heft M 52). Bremerhaven: Wirtschaftsverlag NW.
- Mourant, R. R. & Rockwell, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors*, 14(4), 325-335.
- Münzer, S., Seufert, T. & Brünken, R. (2009). Learning from multimedia presentations: Facilitation function of animations and spatial abilities. *Learning and Individual Differences*, 19, 481-485.
- Mynttinen, S. (2010). *Finnish novice drivers' competences - compared to the Swedish, Dutch and Austrian novices*. Helsinki: Finnish Transport Safety Agency.
- Nagayama, Y. (1978). Role of visual perception in driving. *IATSS research*, 2, 64-73.
- Oerter, R. (1997). Beiläufiges Lernen - nur eine beiläufige Angelegenheit? In H. Gruber & A. Renkl (Hrsg.), *Wege zum Können. Determinanten des Kompetenzerwerbs*. Bern: Huber.
- Opwis, K., Gold, A., Gruber, H. & Schneider, W. (1990). Zum Einfluss von Expertise auf Gedächtnisleistungen sowie deren Selbsteinschätzung bei Kindern und Erwachsenen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 22, 207-224.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429-434.
- Paas, F. G. W. C. & Van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, 35(4), 737-743.
- Park, B. (2010). Testing the Additivity Hypothesis of Cognitive Load Theory. Saarbrücken: Dissertation, Digitale Bibliothek der Universität des Saarlandes, http://scidok.sulb.uni-saarland.de/volltexte/2010/3478/pdf/Dissertation_BP_2010.pdf.
- Pelz, D. C. & Krupat, E. (1974). Caution profile and driving record of undergraduate males. *Accident Analysis and Prevention*, 6, 45-58.

-
- Petzoldt, T., Weiß, T., Franke, T., Krems, J. F. & Bannert, M. (2011). *Unterstützung der Fahrausbildung durch Lernsoftware* (No. M 219): Berichte der Bundesanstalt für Straßenwesen.
- Pollatsek, A., Narayanaan, V., Pradhan, A. K. & Fisher, D. L. (2006). Using eye movements to evaluate a pc-based risk awareness and perception training program on a driving simulator. *Human Factors*, 48, 447-464.
- Posner, M. I. (1988). Introduction: What is it to be an expert? In M. T. H. Chi, R. Glaser & M. J. Farr (Hrsg.), *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Pradhan, A. K., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A. & Fisher, D. L. (2005). The use of eye movements to evaluate the effects of driver age on risk perception in an advanced driving simulator. *Human Factors*, 47, 840-852.
- Quimby, A. R., Maycock, G., Carter, I. D., Dixon, R. & Wall, J. G. (1986). *Perceptual abilities of accident involved drivers* (TRL Research Report 27). Crowthorne, Berkshire TRL Limited.
- Quimby, A. R. & Watts, G. R. (1981). *Human factors and driving performance* (TRL Report LR1004). Crowthorne, Berkshire: Transport and Road Research Laboratory.
- Rasmussen, J. (1986). *Information processing and human-machine interaction. An approach to cognitive engineering*. New York: Academic Press.
- Rauch, N., Gradenegger, B. & Krüger, H.-P. (2009). Darf ich oder darf ich nicht? Situationsbewusstsein im Umgang mit Nebenaufgaben während der Fahrt. *Zeitschrift für Arbeitswissenschaft*, 63(1), 3-15.
- Regan, M. A., Triggs, T. & Deery, H. (1998). *Training cognitive driving skills: A simulator study*. Paper presented at the 34th Annual Conference of the Ergonomics Society of Australia, Melbourne, Australia.
- Reimann, P. & Chi, M. T. H. (1989). Human expertise. In K. J. Gilhooly (Hrsg.), *Human and machine problem solving* (S. 161-191). New York: Plenum.
- Reimann, P. & Rapp, A. (2008). Expertiserwerb. In A. Renkl (Hrsg.), *Lehrbuch Pädagogische Psychologie* (S. 155-203). Bern: Huber.
- Rikers, R. M. J. P. & Paas, F. (2005). Introduction: Recent advances in expertise research. *Applied Cognitive Psychology*, pp. 145-149.
- Rikers, R. M. J. P., Schmidt, H. G. & Moulart, V. (2005). Biomedical Knowledge: Encapsulated or Two Worlds Apart? *Applied Cognitive Psychology*, 19, 223-231.
- Rumar, K. (1990). The basic driver error: late detection. *Ergonomics*, 33, 1281-1290.
- Sagberg, F. (1998). *Month-by-month changes in accident risk among young novice drivers*. Oslo, Norway: Institute of Transport Economics, Norwegian Centre for Transport Research.
- Sagberg, F. (2002). Driver education from the age of 16: Potential of an extended learning period and increased driving experience to reduce the crash risk of novice drivers. In BASt (Hrsg.), *Zweite Internationale Konferenz "Junge Fahrer und Fahrerinnen"* (Bd. Berichte der BASt, Heft M 143, S. 131-135). Bremerhaven: Wirtschaftsverlag NW.
- Sagberg, F. & Bjornskau, T. (2006). Hazard perception and driving experience among novice drivers. *Accident Analysis and Prevention*, 38, 407-414.

-
- Salden, R. J. C. M., Paas, F., Broers, N. J. & Van Merriënboer, J. J. G. (2004). Mental Effort and Performance as Determinants for the Dynamic Selection of Learning Tasks in Air Traffic Control Training. *Instructional Science*, 32(1-2), 153-172.
- Saß, H., Wittchen, H.-U., Zaudig, M. & Houben, I. (2003). *Diagnostisches und statistisches Manual psychischer Störungen - Textrevision - DSM-IV-TR* Göttingen: Hogrefe.
- Schade, F.-D. (2001). Verkehrsauffälligkeit mit und ohne Unfällen bei Fahranfängern (Reanalyse von Rohdaten der Untersuchung von Hansjosten & Schade, 1997: Legalbewährung von Fahranfängern; Berichte der BASt, Heft M71): Flensburg: Kraftfahrtbundesamt.
- Schade, F.-D. (2002). Einwirkungsmöglichkeiten auf das Verkehrsverhalten von jungen Fahrern mit Hilfe des Verkehrszentralregisters - Bisherige Erkenntnisse und künftige Gestaltungsmöglichkeiten. In BASt (Hrsg.), *Zweite Internationale Konferenz "Junge Fahrer und Fahrerinnen"* (Bd. Berichte der BASt, Heft M 143, S. 149-152). Bremerhaven: Wirtschaftsverlag NW.
- Schank, R. C. (1982). *Dynamic memory*. Cambridge: Cambridge University Press.
- Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.
- Schmidt, H. G. & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, 21, 338-351.
- Schmidt, L. (1986). *VIP. Verkehrsspezifischer Itempool. Testhandbuch*. Wien: Kuratorium für Verkehrssicherheit.
- Schnotz, W., Böckheler, J. & Grondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education*, 14, 245-265.
- Schwalm, M. (2008). *Pupillometrie als Methode zur Erfassung mentaler Beanspruchungen im automotiven Kontext*. Dissertation, Universität des Saarlandes, Saarbrücken.
- Shiffrin, R. M. & Atkinson, R. C. (1971). The control of short-term memory. *Scientific American*, 225, 82-90.
- Simon, H. A. & Chase, W. G. (1973). Skill in chess. *American Scientist*, 61, 394-403.
- Smith, S. S., Horswill, M. S., Chambers, B. & Wetton, M. A. (2009a). Hazard perception in novice and experienced drivers: The effects of sleepiness. *Accident Analysis & Prevention*, 41(4), 729-733.
- Smith, S. S., Horswill, M. S., Chambers, B. & Wetton, M. A. (2009b). *Sleepiness and hazard perception while driving* (No. RSRG 2009-001). Canberra, Australia: Department of Infrastructure, Transport, Regional Development and Local Government.
- Soliday, S. M. (1974). Relationship between age and hazard perception in automobile drivers. *Perceptual and Motor Skills*, 39, 335-338.
- Solso, R. L. (2005). *Kognitive Psychologie* (M. Reiss, Trans.). Heidelberg: Springer Medizin Verlag.
- Spicer, R. A. (1964). *Human factors in traffic accidents. Final Report*. (No. AC-00055. USPHS Research Grant, May 1964). Honolulu, Hawaii: Department of Health.

-
- Spiro, R. J., Coulson, R. L., Feltovich, R. J. & Anderson, D. K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In V. Patel (Hrsg.), *Tenth annual conference of the Cognitive Science Society* (S. 375-385). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J. & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 31(5), 24-33.
- Spiro, R. J. & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Hrsg.), *Cognition, education, and multimedia: Exploring ideas in high technology* (S. 163-205). Hillsdale, NJ: Erlbaum.
- Spolander, K. (1982). *Inexperienced drivers` behaviour, abilities and attitudes*. Stockholm, Sweden: Swedish National Road Traffic Research Institute.
- Statistisches Bundesamt. (2006). *Unfälle von 18-24-Jährigen im Straßenverkehr*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (2008). *Verkehrsunfälle. Unfälle von 18- bis 24-jährigen im Straßenverkehr 2007*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (2010a). *Verkehrsunfälle. Unfälle von 18-24-Jährigen im Straßenverkehr*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (2010b). *Verkehrsunfälle. Unfallentwicklung im Straßenverkehr*. Wiesbaden: Statistisches Bundesamt.
- Stein, A. C. & Allen, W. A. (1987, September 9-12, 1986). *The effects of alcohol on driver decision making and risk taking*. Paper presented at the Alcohol, drugs and traffic safety - T86. Proceedings of the Tenth International Conference on Alcohol, Drugs and Traffic Safety, Amsterdam.
- Stelling, D. (2005). *Psychologische Faktoren und Situationsbewusstsein* (No. 14). Hamburg: Deutsches Zentrum für Luft- und Raumfahrt Institut für Luft- und Raumfahrtmedizin, Luft- und Raumfahrtpsychologie.
- Sternberg, R. J. (1990). *Metaphors of mind: conceptions of the nature of human intelligence*. Cambridge, UK: Cambridge University Press.
- Stiensmeier-Pelster, J. (2007). *Abschlussbericht zum Niedersächsischen Modellversuch Begleitetes Fahren ab 17*: Niedersächsisches Ministerium für Wirtschaft, Arbeit und Verkehr.
- Stiensmeier-Pelster, J. & Schöne, C. (2005). *Vorläufige Ergebnisse der wissenschaftlichen Evaluation*. Paper presented at the Symposium zum Niedersächsischen Modellversuch "Begleitetes Fahren mit 17", Berlin.
- Sturzbecher, D. (2008). Methodische Grundlagen der praktischen Fahrerlaubnisprüfung. In D. Sturzbecher, J. Bönninger & M. Rüdell (Hrsg.), *Optimierung der praktischen Fahrerlaubnisprüfung* (S. 16-42). Dresden: TÜV | DEKRA arge tp 21.
- Sturzbecher, D., Kasper, D., Bönninger, J. & Rüdell, M. (2008). *Evaluation der theoretischen Fahrerlaubnisprüfung - Methodische Konzeption und Ergebnisse des Revisionsprojekts*. Dresden: TÜV | Dekra arge tp 21.

-
- Sümer, N., Berfu, Ü. A. & Birdal, A. (2007). *Assessment of Hazard Perception latencies using real life and animated traffic hazards: comparison of novice and experienced drivers*. Paper presented at the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Stevenson, Washington.
- Summala, H. & Näätänen, R. (1988). The zero-risk theory and overtaking decisions. In T. Rothengatter & R. A. d. Bruin (Hrsg.), *Road user behaviour: Theory and research* (S. 82-92). Assen: Van Gorcum.
- Svenson, O. (2009). Driving speed changes and subjective estimates of time savings, accident risks and braking. *Applied Cognitive Psychology*, 23, 543-560.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Sweller, J. & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 186-233.
- Swets, J. A. (1961). Is there a sensory threshold? *Science*, 134, 168-177.
- Thomas, J. R., Gallgher, J. & Lowry, K. (2003). *Developing motor and sport expertise: Meta-analytic findings*. Paper presented at the North American Society for the Psychology of Sport and Physical Activity, 5-7 June, Savannah, GA.
- Tversky, B., Morrison, J.-B. & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human Computer Studies*, 57, 247-262.
- Underwood, G. (2007). Visual attention and the transition from novice to advanced drivers. *Ergonomics*, 50(8), 1235-1249.
- Underwood, G., Crundall, D. & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, In Press, Corrected Proof.
- van Gog, T., Ericsson, K. A., Rikers, R. M. J. P. & Paas, F. (2005). Instructional design for advanced learners: establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development*, 53(3), 73-81.
- Vollrath, M. (2010). Welche Fehler führen zu Unfällen? *Zeitschrift für Verkehrssicherheit*, 56(1), 31-36.
- Ward, P. & Williams, A. M. (2003). Perceptual and cognitive skill development in soccer: The multidimensional nature of expert performance. *Journal of Sport and Exercise Psychology*, 25, 93-111.
- Ward, P., Williams, A. M. & Hancock, P. A. (2006). Simulation for performance and training. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Hrsg.), *The cambridge handbook of expertise and expert performance* (S. 243-262). New York: Cambridge University Press.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45-66). Seattle: Hogrefe & Huber.
- Weiß, T., Petzold, T., Bannert, M. & Krems, J. F. (2007). *Einsatz von computergestützten Medien und Fahrsimulatoren in Fahrausbildung, Fahrerweiterbildung und Fahrerlaubnisprüfung*: Bundesanstalt für Straßenwesen.

-
- Wells, P., Tong, S., Sexton, B., Grayson, G. & Jones, E. (2008). *Cohort II: A study of learner and new drivers. Vol. 1 - Main Report. Road Safety Research Report* (No. 81). London:
Department for Transport.
- Wetton, M. A., Horswill, M. S., Hatherly, C., Wood, J. M., Pachana, N. A. & Anstey, K. J. (2010). The development and validation of two complementary measures of drivers' hazard perception ability. *Accident Analysis and Prevention*, 42, 1232-1239.
- Williams, A. F. (2003). Teenage drivers: patterns of risk. *Journal of Safety Research*, 34, 5-15.
- Williams, A. F. & O'Neill, B. (1974). On-the-road driving records of licensed race drivers. *Accident Analysis & Prevention*, 6(3-4), 263-270.
- Williams, A. M. & Ericsson, K. A. (2005). Perceptual-cognitive expertise in sport: Some considerations when applying the expert performance approach. *Human Movement Science*, 24(3), 283-307.
- Willmes-Lenz, G. (2010). Unfallrisiko "Junge Fahrer" - neue Lösungsansätze. Arbeitskreis VII. Unfallrisiko "Junge Fahrer". In *Deutscher Verkehrsgerichtstag - Deutsche Akademie für Verkehrswissenschaften e.V. -Tagungsband zum 48. Deutschen Verkehrsgerichtstag 2010* (Bd. Köln): Luchterhand.
- Wright, P., Milroy, R. & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education*, 14(2), 203-224.
- Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Tabellenverzeichnis

Tabelle 2.1 <i>Transformation mentaler Funktionen auf den Stufen des Modells des Fertigkeitserwerbs nach Dreyfus und Dreyfus (1980)</i>	20
Tabelle 4.1 <i>Verteilung der Versuchsteilnehmer auf die Bedingungen (Vorstudie 1)</i>	82
Tabelle 4.2 <i>Variation der Parameter Geschwindigkeit, Abstand und Witterung im Aufgabenblock „Anhalteweg“</i>	88
Tabelle 4.3 <i>Variation der Parameter Geschwindigkeit, Abstand und Witterung im Aufgabenblock „Sicherheitsabstand“</i>	90
Tabelle 4.4 <i>Variation der Parameter Geschwindigkeit des Ego-Fahrzeuges (PKW), Geschwindigkeit des überholten Fahrzeuges (LKW), Abstand und Witterung im Aufgabenblock „Sicherheitsabstand“</i>	91
Tabelle 4.5 <i>Itemkennwerte der Wissensaufgaben mit Bildmaterial (Übersicht)</i>	97
Tabelle 4.6 <i>Itemkennwerte der Einschätzungsaufgaben (Übersicht)</i>	98
Tabelle 4.7 <i>Mittelwerte und Standardabweichungen der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch; Maximalpunktzahl je Aufgabenart: 22)</i>	100
Tabelle 4.8 <i>Mittelwerte und Standardabweichungen der Testeffizienz der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch)</i>	103
Tabelle 4.9 <i>Korrelationen (Pearson r) zwischen Unfallzahlen/ Kritischen Situationen und der Leistung in den Wissens-/ Einschätzungs-/ Textaufgaben</i>	105
Tabelle 4.10 <i>Verteilung der Versuchsteilnehmer auf die Bedingungen (Vorstudie 2)</i>	117
Tabelle 4.11 <i>Beschreibung der Reaktionszeitaufgaben</i>	122
Tabelle 4.12 <i>Itemkennwerte der Reaktionszeitaufgaben (Übersicht)</i>	128
Tabelle 4.13 <i>Ergebnisse in der Kontrollvariablen Sensationssuche: Mittelwerte und Standardabweichungen in den TAS-Skalen</i>	129
Tabelle 4.14 <i>Ergebnisse in den Reaktionszeitaufgaben (mittlere Lösungswahrscheinlichkeiten): Mittelwerte und Standardabweichungen der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch)</i>	131

Tabelle 4.15 <i>Reaktionszeitaufgaben (z-Werte): Mittelwerte und Standardabweichungen der Experten und Novizen in den verschiedenen Bedingungen (statisch vs. dynamisch)</i>	134
Tabelle 5.1 <i>Stichprobe getrennt nach Geschlecht und Expertise zu den einzelnen Messzeitpunkten (N = 154)</i>	154
Tabelle 5.2 <i>Häufigkeitsverteilungen (%) der Experten und Novizen auf die unterschiedlichen Paralleltestversionen an den unterschiedlichen Messzeitpunkten</i>	154
Tabelle 5.3 <i>Einwohnerzahlen der Wohnorte der Versuchspersonen (%)</i>	155
Tabelle 5.4 <i>Höchster Bildungsabschluss der Versuchspersonen (%).....</i>	156
Tabelle 5.5 <i>Berufliche Situation der Versuchspersonen (%)</i>	156
Tabelle 5.6 <i>Absolvierte Praktische Fahrstunden (%)</i>	162
Tabelle 5.7 <i>Fahrerfahrung der Teilnehmer in km (%).....</i>	163
Tabelle 5.8 <i>Durchschnittliche wöchentliche Fahrleistung der Teilnehmer (%).....</i>	164
Tabelle 5.9 <i>Durchschnittliche Anzahl von selbstverschuldeten Unfällen, Verwarnungsgeldern und kritischen Situationen (Standardabweichung)</i>	165
Tabelle 5.10 <i>Parallelität der Testformen A, B und C: Mittelwerte (Standardabweichung) der Anzahl der richtigen Lösungen in den Wissens- und Reaktionszeitaufgaben</i>	166
Tabelle 5.11 <i>Korrelative Zusammenhänge (Pearson r) zwischen den Leistungen im Paper Folding Test, Hidden Patterns Test und den Leistungen in den verkehrsspezifischen Leistungsmaßen.....</i>	168
Tabelle 5.12 <i>Korrelative Zusammenhänge (Pearson r) zwischen den Ergebnissen in der Sensation-Seeking-Skala und den Leistungen in den verkehrsspezifischen Leistungsmaßen</i>	169
Tabelle 5.13 <i>Mittelwerte und Standardabweichungen der Experten und Novizen in den Wissens-, Regel- und Reaktionszeitaufgaben zum ersten Messzeitpunkt</i>	170
Tabelle 5.14 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen zum ersten Messzeitpunkt</i>	171
Tabelle 5.15 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: Reaktionszeiten) getrennt nach Aufgabentypen zum ersten Messzeitpunkt</i>	172

Tabelle 5.16 <i>Mittelwerte und Standardabweichungen der Experten und Novizen in den Wissens-, Regel- und Reaktionszeitaufgaben zum zweiten Messzeitpunkt</i>	174
Tabelle 5.17 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen zum zweiten Messzeitpunkt</i>	175
Tabelle 5.18 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: Reaktionszeiten) getrennt nach Aufgabentypen zum zweiten Messzeitpunkt</i>	176
Tabelle 5.19 <i>Mittelwerte und Standardabweichungen der Experten und Novizen in den Wissens-, Regel- und Reaktionszeitaufgaben zum dritten Messzeitpunkt</i>	177
Tabelle 5.20 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen zum dritten Messzeitpunkt</i>	178
Tabelle 5.21 <i>Zusammenhänge zwischen der Visualisierungsfähigkeit und der Leistung in den verschiedenen Aufgabentypen des Reaktionstests (N = 120)</i>	179
Tabelle 5.22 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: Reaktionszeiten) getrennt nach Aufgabentypen zum dritten Messzeitpunkt</i>	180
Tabelle 5.23 <i>Deskriptive Ergebnisse der Längsschnittuntersuchung (N = 100) (vgl. Malone et al., 2012)</i>	183
Tabelle 5.24 <i>Ergebnisse der Kovarianzanalyse für die Wissensaufgaben</i>	184
Tabelle 5.25 <i>Ergebnisse der Kovarianzanalyse für die Regelaufgaben</i>	186
Tabelle 5.26 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen</i>	188
Tabelle 5.27 <i>Ergebnisse der Kovarianzanalyse für den Reaktionstest (AV korrekte Lösungen)</i>	189
Tabelle 5.28 <i>Mittelwerte und Standardabweichungen der Experten und Novizen im Reaktionstest (AV: korrekte Lösungen) getrennt nach Aufgabentypen</i>	192
Tabelle 5.29 <i>Ergebnisse der Kovarianzanalyse für den Reaktionstest (AV Reaktionszeiten)</i> 193	

Abbildungsverzeichnis

Abbildung 2.1 <i>Modell der Informationsverarbeitung (nach Atkinson & Shiffrin, 1969; Shiffrin & Atkinson, 1971, aus Leutner & Brünken, 2002, S.77).....</i>	16
Abbildung 2.2 <i>Unfälle mit Delikteintragung im Verkehrszentralregister von Fahranfängern in den ersten vier Jahren ihrer Fahrerkarriere (Jahr des Fahrerlaubniserwerbs: 1987; Männer: N=5205; Frauen: N=6095; aus Schade, 2001, S. 3).....</i>	33
Abbildung 2.3 <i>Darstellung von Anfänger- und Jugendlichkeitsrisiko anhand der vorhergesagten jährlichen Unfallraten für Fahranfänger mit unterschiedlichem Einstiegsalter (nach Maycock, Lockwood, & Lester, 1991; aus Leutner, Brünken, & Willmes-Lenz, 2009, S.2).....</i>	38
Abbildung 2.4 <i>Hierarchisches Modell des Fahrverhaltens (nach Hatakka et al., 1999).....</i>	41
Abbildung 2.5 <i>Vereinfachtes Modell der Handlungskontrolle auf drei Ebenen (übersetzt von Leutner & Brünken, 2002; nach Rasmussen, 1986).....</i>	42
Abbildung 2.6 <i>Modell des Fahrverhaltens als Reaktion auf eine potenzielle Gefahr (nach Deery, 1999).....</i>	48
Abbildung 2.7 <i>Modell des Risikoverhaltens (nach Grayson et al., 2003).....</i>	49
Abbildung 3.1 <i>Veranschaulichung der empirischen Vorgehensweise.....</i>	77
Abbildung 3.2 <i>Veranschaulichung der Hypothesen für die Längsschnittstudie.....</i>	77
Abbildung 4.1 <i>Originalbild der Aufgabe 2.1.08-008 (links) und mit dem VICOM-Editor erstelltes Bild (rechts)(aus Malone, Biermann, Buch & Brünken, 2012; S. 36).....</i>	85
Abbildung 4.2 <i>Endbild (Screenshot) einer Beispielaufgabe aus der dynamischen Version des Testmaterials.....</i>	86
Abbildung 4.3 <i>Statisches Bild der Einschätzungsaufgabe 1 aus dem Aufgabenblock „Anhalteweg“ (aus Malone et al., 2012; S. 37).....</i>	88
Abbildung 4.4 <i>Statisches Bild der Einschätzungsaufgabe 2 aus dem Aufgabenblock „Anhalteweg“ (aus Malone et al., 2012; S. 37).....</i>	89
Abbildung 4.5 <i>Statisches Bild der Einschätzungsaufgabe 12 aus dem Aufgabenblock „Sicherheitsabstand“.....</i>	90

Abbildung 4.6 <i>Statisches Bild der Einschätzungsaufgabe 20 aus dem Aufgabenblock „Überholweg“</i>	92
Abbildung 4.7 <i>Einschätzungsaufgabe 1 in der statischen Version mit Aufgabenstellung und Einschätzung der Sicherheit</i>	92
Abbildung 4.8 <i>Mittlere Leistungen von Experten und Novizen in den Wissensaufgaben getrennt für die beiden Präsentationsformen statisch und dynamisch (Maximalpunktzahl: 22, * $p < .05$)(vgl. Malone et al., 2012)</i>	101
Abbildung 4.9 <i>Statische Version der Reaktionszeitaufgabe Nr. 28 (Hazard Perception) (aus Malone et al., 2012; S. 45)</i>	120
Abbildung 4.10 <i>Statische Version der Reaktionszeitaufgabe Nr. 15 (Vorschriftaufgabe)</i>	121
Abbildung 4.11 <i>Ergebnisse zu Reaktionszeitaufgaben (Lösungswahrscheinlichkeit): Haupteffekt Aufgabentyp</i>	132
Abbildung 4.12 <i>Ergebnisse zu Reaktionszeitaufgaben (Lösungswahrscheinlichkeit): Interaktion Aufgabentyp*Präsentationsform</i>	132
Abbildung 4.13 <i>Ergebnisse zu Reaktionszeitaufgaben (z-Werte): Interaktion Aufgabentyp*Präsentationsform</i>	135
Abbildung 4.14 <i>Ergebnisse zu Reaktionszeitaufgaben (z-Werte): Interaktion Aufgabentyp*Präsentationsform*Expertise</i>	136
Abbildung 5.1 <i>Regelaufgabe Nr. 7</i>	158
Abbildung 5.2 <i>Längsschnittstudie, MZP1: Ergebnisse zu Reaktionszeitaufgaben (z-Werte): Interaktion Aufgabentyp*Expertise</i>	173
Abbildung 5.3 <i>Ergebnisse zu Reaktionszeitaufgaben (Lösungswahrscheinlichkeit) zu Messzeitpunkt 3: Haupteffekt Aufgabentyp</i>	179
Abbildung 5.4 <i>Leistungsentwicklung der Experten und Novizen in den Wissensaufgaben über die drei Messzeitpunkte hinweg (vgl. Malone et al., 2012)</i>	184
Abbildung 5.5 <i>Leistungsentwicklung der Experten und Novizen in den Regelaufgaben über die drei Messzeitpunkte hinweg (vgl. Malone et al., 2012)</i>	186
Abbildung 5.6 <i>Leistungsentwicklung der Experten und Novizen in den Reaktionszeitaufgaben (AV: korrekte Lösungen) über die drei Messzeitpunkte hinweg (vgl. Malone et al., 2012)</i> ...	187

Abbildung 5.7 <i>Leistungsunterschiede zwischen Experten und Novizen in den Reaktionszeitaufgaben (AV: Reaktionszeiten) an drei Messzeitpunkten (vgl. Malone et al., 2012)</i>	191
--	-----

Anhang

Anhang A: Zusätzliche Berechnungen

A1. Vorstudie 1: Itemkennwerte

A2. Vorstudie 2: Itemkennwerte

A3. Längsschnittstudie: Kovarianzanalysen mit alternativen z-Werten

Anhang B: Versuchsmaterial (DVD auf Anfrage bei der Autorin)

Anhang A**A1. Vorstudie 1: Itemkennwerte**

Tabelle A1.1 Itemstatistiken und Reliabilität: Wissensaufgaben mit Bildmaterial (gesamt)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
1_1_02_008	.88	.33	.53	.15	.73
1_1_02_025	.40	.49	.28	-.08	.74
1_1_02_105	.76	.43	.32	.01	.74
1_1_02_110	.31	.46	.36	-.20*	.73
1_1_03_101	.58	.50	.43	.08	.73
1_1_03_103	.59	.49	.38	.08	.73
1_1_04_104	.45	.50	.47	.12	.72
1_1_04_105	.44	.50	.22	-.03	.75
1_1_05_004	.59	.49	.41	-.04	.73
1_1_05_101	.63	.48	.43	.07	.73
1_1_06_107	.78	.41	.30	-.01	.74
1_1_06_110	.68	.47	.47	.23	.73
2_1_06_004	.87	.34	.19	.17	.75
2_1_06_102	.80	.40	.35	.23	.74
2_1_08_003	.39	.49	.14	-.06	.75
2_1_08_006	.95	.22	.09	.09	.75
2_1_08_008	.49	.50	.32	.17	.74
2_1_08_101	.68	.47	.17	-.04	.75
2_2_37_005	.92	.26	.13	.05	.75
2_1_03_017	.79	.41	.15	-.08	.75
2_1_03_024	.75	.44	.10	-.16	.75
2_1_06_006	.97	.18	.33	.20*	.74
M	.67	.42	.30	.04	.74

Alpha: .75

Tabelle A1.2 Itemstatistiken und Reliabilität: Wissensaufgaben mit Bildmaterial (statisch)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
1_1_02_008	.92	.28	.61	.33*	.73
1_1_02_025	.32	.47	.33	-.02	.74
1_1_02_105	.71	.46	.39	.09	.73
1_1_02_110	.29	.46	.27	-.02	.74
1_1_03_101	.59	.50	.44	.14	.73
1_1_03_103	.56	.50	.29	-.06	.74
1_1_04_104	.37	.49	.42	.29*	.73
1_1_04_105	.46	.50	.33	.02	.74
1_1_05_004	.64	.48	.46	.10	.73
1_1_05_101	.61	.49	.47	.17	.72
1_1_06_107	.76	.43	.34	.13	.74
1_1_06_110	.53	.50	.42	.35**	.73
2_1_06_004	.86	.35	.22	.33*	.74
2_1_06_102	.80	.41	.31	.55**	.74
2_1_08_003	.29	.46	.19	-.09	.75
2_1_08_006	.95	.22	-.04	-.06	.75
2_1_08_008	.46	.50	.27	.23	.74
2_1_08_101	.73	.45	.09	-.03	.75
2_2_37_005	.92	.28	.11	-.04	.75
2_1_03_017	.85	.36	.18	.08	.75
2_1_03_024	.71	.46	.06	-.28	.76
2_1_06_006	.95	.22	.42	.25	.74
M	.65	.42	.30	.11	.74

Alpha: .75

Tabelle A1.3 Itemstatistiken und Reliabilität: Wissensaufgaben mit Bildmaterial (dynamisch)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
1_1_02_008	.84	.37	.50	.01	.73
1_1_02_025	.48	.50	.16	-.11	.75
1_1_02_105	.80	.40	.30	-.07	.74
1_1_02_110	.33	.47	.46	-.36**	.73
1_1_03_101	.57	.50	.45	.01	.73
1_1_03_103	.62	.49	.46	-.09	.73
1_1_04_104	.52	.50	.48	-.02	.73
1_1_04_105	.43	.50	.12	-.08	.76
1_1_05_004	.54	.50	.40	-.18	.74
1_1_05_101	.66	.48	.40	-.02	.74
1_1_06_107	.80	.40	.26	-.16	.75
1_1_06_110	.82	.39	.53	.14	.73
2_1_06_004	.87	.34	.20	.01	.75
2_1_06_102	.80	.40	.35	-.07	.74
2_1_08_003	.49	.50	.13	-.02	.76
2_1_08_006	.95	.22	.28	.23	.75
2_1_08_008	.52	.50	.33	.11	.74
2_1_08_101	.64	.48	.27	-.06	.75
2_2_37_005	.93	.25	.17	.14	.75
2_1_03_017	.74	.44	.19	-.21	.75
2_1_03_024	.79	.41	.08	-.03	.76
2_1_06_006	.98	.13	.18	.13	.75
M	.69	.42	.30	-.03	.74

Alpha: .75

Tabelle A1.4 Itemstatistiken und Reliabilität Wissensaufgaben Textaufgaben (ohne Bildmaterial)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
1_1_03_003	.66	.48	.36	.09	.59
1_1_03_104	.53	.50	.31	-.01	.61
1_1_04_101	.45	.50	.26	.09	.62
1_1_05_003	.80	.40	.42	.11	.58
1_1_06_103	.78	.42	.22	.25**	.62
1_1_07_111	.85	.36	.25	.21*	.62
2_1_02_001	.44	.50	.02	-.10	.68
2_1_03_108	.38	.49	.39	.06	.59
2_1_04_102	.73	.44	.50	.07	.56
2_1_05_006	.69	.46	.40	.12	.59
M	.63	.46	.31	.09	.61

Alpha: .63

Tabelle A1.5 Itemstatistiken und Reliabilität: Einschätzungsaufgaben (gesamt)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
Aufg_1_Erg	.73	.44	.12	-.05	.25
Aufg_2_Erg	.78	.42	-.11	-.11	.31
Aufg_3_Erg	.81	.40	.09	-.12	.26
Aufg_4_Erg	.39	.49	-.01	-.13	.29
Aufg_5_Erg	.80	.40	-.14	.03	.32
Aufg_6_Erg	.41	.50	.12	.04	.25
Aufg_7_Erg	.63	.48	.08	.11	.26
Aufg_8_Erg	.76	.43	-.01	.09	.29
Aufg_9_Erg	.66	.48	.14	-.09	.24
Aufg_61_Erg	.59	.49	.20	.23	.22
Aufg_62_Erg	.72	.45	.10	.14	.26
Aufg_10_Erg	.62	.49	-.11	-.06	.32
Aufg_12_Erg	.67	.47	-.12	.00	.32
Aufg_112_Erg	.90	.30	.01	-.04	.28
Aufg_13_Erg	.80	.40	.25	-.02	.22
Aufg_14_Erg	.59	.49	.06	-.04	.27
Aufg_15_Erg	.67	.47	.15	.07	.24
Aufg_16_Erg	.68	.47	.14	.03	.24
Aufg_17_Erg	.83	.37	.10	-.02	.26
Aufg_18_Erg	.76	.43	.12	-.03	.25
Aufg_19_Erg	.80	.40	.14	.07	.25
Aufg_20_Erg	.91	.29	.26	.16	.23
M	.77	.43	.07	.01	.27

Alpha .28

Tabelle A1.6 Itemstatistiken und Reliabilität: Einschätzungsaufgaben (statisch)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
Aufg_1_Erg	.69	.46	.16	-.09	.39
Aufg_2_Erg	.68	.47	-.16	-.12	.46
Aufg_3_Erg	.88	.33	.26	-.02	.38
Aufg_4_Erg	.39	.49	-.02	-.10	.43
Aufg_5_Erg	.83	.38	-.15	-.05	.45
Aufg_6_Erg	.44	.50	.26	.20	.36
Aufg_7_Erg	.63	.49	.25	.35**	.37
Aufg_8_Erg	.66	.48	.19	.21	.38
Aufg_9_Erg	.63	.49	.27	.14	.36
Aufg_61_Erg	.76	.43	-.13	-.03	.45
Aufg_62_Erg	.68	.47	.34	-.05	.35
Aufg_10_Erg	.61	.49	-.13	-.11	.46
Aufg_12_Erg	.71	.46	-.24	.17	.48
Aufg_112_Erg	.88	.33	.15	-.02	.40
Aufg_13_Erg	.75	.44	.34	.17	.35
Aufg_14_Erg	.61	.49	-.01	-.04	.43
Aufg_15_Erg	.73	.45	.21	-.03	.38
Aufg_16_Erg	.73	.45	.15	.05	.39
Aufg_17_Erg	.86	.35	.20	-.07	.39
Aufg_18_Erg	.73	.45	.27	-.03	.37
Aufg_19_Erg	.75	.44	.18	.01	.39
Aufg_20_Erg	.90	.31	.28	.14	.38
M	.71	.44	.12	.02	.40

Alpha: .41

Tabelle A1.7 Itemstatistiken und Reliabilität: Einschätzungsaufgaben (dynamisch)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
Aufg_1_Erg	.77	.42	.07	.01	.06
Aufg_2_Erg	.87	.34	.0	.01	.09
Aufg_3_Erg	.74	.44	-.04	-.09	.11
Aufg_4_Erg	.39	.49	.02	-.21	.08
Aufg_5_Erg	.77	.42	-.13	-.15	.14
Aufg_6_Erg	.38	.49	-.04	.09	.11
Aufg_7_Erg	.56	.50	.15	-.11	.02
Aufg_8_Erg	.77	.42	.0	.11	.09
Aufg_9_Erg	.64	.48	-.12	.09	.14
Aufg_61_Erg	.75	.43	.13	.08	.04
Aufg_62_Erg	.64	.48	-.09	.20	.13
Aufg_10_Erg	.62	.49	-.09	-.12	.13
Aufg_12_Erg	.62	.49	.01	-.02	.09
Aufg_112_Erg	.92	.28	-.17	-.16	.13
Aufg_13_Erg	.85	.36	.13	-.06	.05
Aufg_14_Erg	.57	.50	.14	-.22	.02
Aufg_15_Erg	.61	.49	.1	-.05	.04
Aufg_16_Erg	.64	.48	.14	.15	.02
Aufg_17_Erg	.80	.40	.00	.01	.09
Aufg_18_Erg	.79	.41	-.06	.01	.11
Aufg_19_Erg	.85	.36	.10	-.03	.05
Aufg_20_Erg	.92	.28	.23	.15	.03
M	.73	.43	.02	-.01	.08

Alpha .09

A2. Vorstudie 2: Itemkennwerte

Tabelle A2.1 Itemstatistiken und Reliabilität: Reaktionszeitaufgaben (AV: korrekte Lösungen, gesamt)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
2	.58	.50	.49	-.11	.75
3	.77	.42	.51	-.06	.75
4	.90	.31	.24	.14	.77
5	.84	.37	.08	.24**	.77
6	.87	.34	.05	.11	.77
7	.81	.40	.44	.17*	.76
8	.83	.38	.36	.13	.76
9	.63	.49	-.02	.09	.78
10	.75	.44	.30	.07	.76
11	.51	.50	.24	.13	.77
12	.87	.33	.41	.07	.76
13	.51	.50	.21	.00	.77
14	.55	.50	.43	.11	.76
15	.22	.41	.30	.15	.76
16	.75	.44	.32	.14	.76
17	.75	.43	.20	.12	.77
18	.50	.50	-.03	.20*	.78
19	.75	.43	.29	.12	.76
20	.24	.43	.03	-.07	.78
21	.50	.50	.03	.36**	.78
22	.84	.37	.22	.23**	.77
23	.64	.48	.35	.29**	.76
24	.62	.49	.42	.03	.76
25	.64	.48	.40	.03	.76
26	.54	.50	.50	.10	.75
27	.56	.50	.39	.03	.76
28	.33	.47	.18	.10	.77
29	.85	.36	.17	.13	.77
30	.51	.50	.57	.06	.75
31	.51	.50	.13	.10	.77
32	.76	.43	.38	.05	.76
M	.64	.44	.28	.28	.77

Alpha .77

Tabelle A2.2 Itemstatistiken und Reliabilität: Reaktionszeitaufgaben (AV korrekte Lösungen, statisch)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
2	.87	.34	.35	.02	.72
3	.93	.25	.41	.27*	.72
4	.98	.13	-.06	.13	.73
5	.84	.37	-.14	.20	.75
6	.84	.37	.04	.15	.74
7	.95	.22	.42	.24	.72
8	.92	.28	.36	.23	.72
9	.46	.50	.31	.15	.72
10	.92	.28	.44	.31*	.72
11	.62	.49	.39	.07	.71
12	.98	.13	.45	.13	.72
13	.72	.45	.27	.07	.72
14	.90	.30	.31	.23	.72
15	.28	.45	.32	.04	.72
16	.90	.30	-.07	.12	.74
17	.87	.34	.24	.11	.72
18	.38	.49	.31	.08	.72
19	.90	.30	.26	.23	.72
20	.26	.44	.02	.10	.74
21	.39	.49	.22	.32*	.73
22	.90	.30	.00	.23	.74
23	.77	.42	.15	.51**	.73
24	.84	.37	.37	.23	.72
25	.80	.40	.42	.02	.71
26	.85	.36	.48	.37**	.71
27	.69	.47	.30	.00	.72
28	.43	.50	.02	.20	.74
29	.79	.41	.45	.14	.71
30	.84	.37	.37	.20	.72
31	.56	.50	.06	.20	.74
32	.87	.34	.51	.25	.71
M	.75	.37	.26	.18	.72

Alpha .72

Tabelle A2.3 Itemstatistiken und Reliabilität: Reaktionszeitaufgaben (richtige Lösungen, dynamisch)

Nr.	Schwierigkeit	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
2	.34	.48	.20	-.16	.58
3	.63	.49	.41	-.18	.55
4	.82	.39	.16	.19	.58
5	.85	.36	.34	.27*	.57
6	.89	.32	.22	.06	.58
7	.68	.47	.30	.21	.57
8	.75	.43	.26	.10	.57
9	.77	.43	.21	.00	.58
10	.60	.49	-.05	.01	.61
11	.42	.50	-.03	.21	.61
12	.78	.42	.30	.10	.57
13	.33	.47	-.26	.01	.63
14	.26	.44	-.02	.18	.60
15	.16	.37	.23	.30*	.58
16	.62	.49	.30	.20	.57
17	.66	.48	-.04	.16	.61
18	.60	.49	.03	.29*	.60
19	.63	.49	.06	.11	.60
20	.22	.42	-.03	.03	.60
21	.59	.50	.20	.38**	.58
22	.79	.41	.27	.26*	.57
23	.53	.50	.35	.17	.56
24	.44	.50	.16	.05	.58
25	.51	.50	.19	.07	.58
26	.27	.45	.09	.01	.59
27	.45	.50	.34	.09	.56
28	.25	.43	.13	.03	.59
29	.90	.30	.27	.11	.58
30	.23	.43	.31	.07	.57
31	.48	.50	.14	.02	.59
32	.67	.47	.20	.05	.58
M	.55	.45	.17	.11	.58

Alpha .58

Tabelle A2.4 Itemstatistiken und Reliabilität: Reaktionszeitaufgaben (AV: Reaktionszeit, gesamt)

Nr.	M	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
2	0	1	.62	-.04	.80
3	0	1	.53	-.16	.81
4	0	1	.30	-.13	.82
9	0	1	.23	-.24	.82
10	0	1	.24	-.16	.82
11	0	1	.24	.02	.82
12	0	1	.40	-.22	.81
13	0	1	.13	-.18	.83
14	0	1	.64	.24	.80
15	0	1	.65	.24	.80
17	0	1	.37	-.14	.82
18	0	1	-.05	.05	.83
19	0	1	.40	-.10	.82
20	0	1	.36	.16	.82
21	0	1	.21	-.20	.82
23	0	1	.31	.03	.82
24	0	1	.45	-.18	.81
25	0	1	.42	-.10	.81
26	0	1	.66	-.04	.80
27	0	1	.25	-.18	.82
28	0	1	.28	-.03	.82
29	0	1	.13	-.18	.83
30	0	1	.58	-.17	.81
31	0	1	.24	-.03	.82
32	0	1	.41	-.18	.81
M	0	1	.36	-.08	.82

Alpha .82

Tabelle A2.5 Itemstatistiken und Reliabilität: Reaktionszeitaufgaben (AV: Reaktionszeit, statisch)

Nr.	M	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
2	-.75	.20	.34	-.02	.69
3	-.40	.45	.52	-.16	.67
4	-.30	.09	.40	-.12	.69
9	-.13	.35	.35	-.14	.68
10	-.29	.68	.42	-.20	.67
11	-.30	.81	.26	.16	.68
12	-.36	.13	.46	-.28	.69
13	-.16	1.27	.26	-.19	.69
14	-.86	.21	.41	-.02	.69
15	-.74	.44	.34	.31	.68
17	-.49	.33	.50	-.17	.68
18	.10	1.04	.20	.17	.70
19	-.43	.44	.33	-.06	.68
20	-.51	.55	.05	.26	.70
21	-.08	.83	.18	-.09	.69
23	-.21	.85	.09	.05	.70
24	-.48	.60	.40	-.21	.67
25	-.37	.60	.30	-.07	.68
26	-.87	.23	.56	-.10	.68
27	-.09	1.02	.30	-.17	.68
28	-.35	.75	.00	.12	.71
29	.05	.69	.35	-.19	.68
30	-.58	.76	.42	-.16	.67
31	-.25	.61	-.07	-.02	.71
32	-.39	.34	.53	-.14	.68
M	-.37	.57	.32	-.06	.69

Alpha .69

Tabelle A2.6 Itemstatistiken und Reliabilität: Reaktionszeitaufgaben (AV: Reaktionszeit, dynamisch)

Nr.	M	SD	Trennschärfe Korrigierte Item-Skala- Korrelation	Fremdtrennschärfe	Alpha, wenn Item weggelassen
2	.64	.97	.18	-.07	.47
3	.35	1.20	.41	-.23	.43
4	.26	1.31	.15	-.11	.48
9	.11	1.32	.22	-.42	.46
10	.25	1.16	-.08	-.05	.52
11	.26	1.08	-.08	-.27	.51
12	.31	1.28	.26	-.15	.46
13	.13	.67	-.23	-.28	.52
14	.74	.79	.01	.67	.50
15	.64	.91	.22	.13	.47
17	.42	1.18	-.06	-.24	.51
18	-.09	.96	-.05	-.23	.51
19	.37	1.18	.11	-.32	.49
20	.44	1.09	-.05	.23	.51
21	.07	1.13	.28	-.41	.45
23	.18	1.09	.34	-.22	.45
24	.41	1.09	.10	-.24	.49
25	.32	1.16	.25	-.13	.46
26	.75	.77	.03	.11	.49
27	.07	.98	.34	-.66	.45
28	.30	1.09	.05	-.20	.49
29	-.05	1.21	.26	-.16	.46
30	.50	.91	.24	-.58	.47
31	.21	1.21	.15	-.25	.48
32	.33	1.24	.19	-.32	.47
M	.32	1.08	.13	-.18	.48

Alpha .49

A3. Längsschnittstudie: Kovarianzanalysen mit alternativen z-Werten

Tabelle A3.1 Ergebnisse der Kovarianzanalyse (AV: Reaktionszeiten). Die Z-Standardisierung erfolgte für jede Aufgabe am Gesamtmittelwert über alle drei Messzeitpunkte hinweg.

Quelle der Varianz	<i>df</i>	<i>F</i>	<i>p</i>
Gruppenvergleiche			
Expertise (E)	1	.36	.55
Visualisierungsfähigkeit (V)	1	1.71	.20
FWG (F)	1	.01	.92
Fehler	76	(.48)	
Messwiederholungsvergleiche ^{a)}			
Zeit (Z)	2	.69	.50
Z x E	2	1.36	.26
Z x V	2	.78	.46
Z x F	2	.77	.47
Fehler (Z)	152	(.65)	
Aufgabentyp (A)	2	<.01	.99
A x E	2	<.01	.95
A x Z	2	.20	.82
A x V	2	.14	.71
A x F	2	.11	.74
A x E x Z	2	3.92	.02
A x Z x V	2	.02	.98
A x Z x F	2	.38	.69
Fehler (A)	76	(.24)	
Fehler (A x Z)	152	(.23)	

Tabelle A3.2 Ergebnisse der Kovarianzanalyse (AV: Reaktionszeiten). Die Z-Standardisierung erfolgte für jede Aufgabe am Gruppenmittelwert der jeweiligen Expertisegruppe.

Quelle der Varianz	<i>df</i>	<i>F</i>	<i>p</i>
Gruppenvergleiche			
Expertise (E)	1	2.77	.10
Visualisierungsfähigkeit (V)	1	.05	.82
FWG (F)	1	2.26	.14
Fehler	76	(.48)	
Messwiederholungsvergleiche ^{a)}			
Zeit (Z)	2	1.53	.22
Z x E	2	.69	.51
Z x V	2	2.56	.08
Z x F	2	.30	.74
Fehler (Z)	152	(.65)	
Aufgabentyp (A)	2	3.00	.09
A x E	2	.27	.61
A x Z	2	2.27	.11
A x V	2	2.08	.15
A x F	2	3.68	.06
A x E x Z	2	.50	.61
A x Z x V	2	.19	.83
A x Z x F	2	2.21	.11
Fehler (A)	76	(.23)	
Fehler (A x Z)	152	(.25)	