

VOM FAHRSIMULATOR BIS ZU NATURALISTIC DRIVING STUDIES

TEIL 2

VOR- UND NACHTEILE VERSCHIEDENER METHODEN DER FAHRVERHALTENSBEOBSACHTUNG

Autor: Dr. Martin Gründl, Ergoneers GmbH

Whitepaper, 03.2016

A decorative graphic consisting of a large, dense cluster of small squares in various shades of blue and white, arranged in a way that suggests a dynamic, expanding shape, possibly representing a data visualization or a stylized representation of a driving environment.

ERGONEERS
FROM SCIENCE TO INNOVATION

ABSTRACT

Immer mehr Fahrerassistenz- und Fahrerinformationssysteme sollen das Autofahren komfortabler, stressfreier oder sicherer machen. Während ihrer Entwicklungsphase und bevor sie Einzug ins Auto finden, durchlaufen sie unterschiedlichste Tests, um zu untersuchen, wie gut der Nutzer die Systeme bedienen kann, ob er mit ihnen das Fahren jederzeit unter Kontrolle hat und wie sich diese Systeme auf sein Fahrverhalten auswirken.

Dazu gibt eine Reihe von Standard-Methoden für von Studien zur Fahrverhaltensbeobachtung: Sie reichen von sogenannten *Naturalistic Driving Studies*, über experimentelle Realfahrzeugstudien, bis hin zum simulierten Fahren in aufwändigen Fahrsimulatoren, in einfachen Sitzkisten oder lediglich der visuellen Darbietung von Verkehrsszenarien am Computer. Die Methoden unterscheiden sich stark im Grad des realitätsnahen Fahrens, der experimentellen Kontrolle und der verwendeten Technik.

In diesem Teil 2 des Whitepapers werden die verschiedenen Methoden der Fahrverhaltensbeobachtung beschrieben und ihre unterschiedlichen Stärken und Schwächen anhand wissenschaftlicher Qualitätskriterien wie Objektivität, Reliabilität sowie interner und externer Validität erläutert. Der für die Verhaltenswissenschaften klassische Zielkonflikt zwischen der Realitätsnähe einer Studie und ihrer methodischen Kontrolle wird dabei ausführlich diskutiert.

Zweck dieses Whitepapers ist es zum einen, dem interessierten Laien ein methodisches Hintergrundwissen zu vermitteln, um die Ergebnisse verkehrspsychologischer Untersuchungen kritisch hinterfragen zu können. Zum anderen soll auch dem Experten und Praktiker ein zusammenfassender Überblick über Vor- und Nachteile verschiedener Untersuchungsmethoden gegeben werden, um für den jeweiligen Untersuchungszweck die geeignetste Methode auswählen zu können.

INHALT

1. WOZU DIE BESCHÄFTIGUNG MIT METHODISCHEN ASPEKTEN BEI FAHRVERHALTENS- STUDIEN?.....	4
2. METHODEN DER FAHRVERHALTENSBEOBSACHTUNG	6
NATURALISTIC DRIVING STUDIES	6
EXPERIMENTELLE REALFAHRZEUGSTUDIEN	9
EXPERIMENTE IM FAHRSIMULATOR.....	11
EINFACHE FORMEN SIMULIERTEN FAHRENS.....	16
3. DER ZIELKONFLIKT ZWISCHEN REALITÄTSNÄHE UND METHODISCHER KONTROLLE....	19
4. FAZIT: EXPERIMENTELLE REALFAHRTSTUDIEN ALS VERNÜNFTIGER KOMPROMISS	22
5. LITERATUR	24

1. WOZU DIE BESCHÄFTIGUNG MIT METHODISCHEN ASPEKTEN BEI FAHRVERHALTENSTUDIEN?

In den letzten Jahren haben immer mehr Assistenz- und Informationssysteme Einzug ins Auto gefunden. Ihr Ziel ist stets, das Autofahren komfortabler, stressfreier oder sicherer zu machen. Fahrzeuge, die beispielsweise automatisch den Abstand zum Vordermann regulieren oder den Fahrer bei der Spurhaltung unterstützen, finden immer mehr Verbreitung. An Navigationssysteme haben wir uns sogar mittlerweile so stark gewöhnt, dass wir sie uns kaum noch aus dem Auto wegdenken können. Auf unbekannten Strecken entlasten sie den Fahrer spürbar von der mental beanspruchenden Navigationsaufgabe und machen dadurch das Fahren sicherer.

Doch all diese Systeme haben nicht nur Vorteile. Sie haben auch ihre Schattenseiten: Das Eingeben eines Ziels in das Navi beispielsweise lenkt den Fahrer extrem vom Verkehr ab und erhöht in dieser Zeitspanne das Unfallrisiko massiv. Auch Assistenzsysteme, die ein teilautomatisiertes Fahren unterstützen, sind nicht ohne Nebenwirkungen, da sich der Fahrer dadurch weniger auf den Verkehr konzentriert und seine Aufmerksamkeit verstärkt Nebentätigkeiten widmet. Die umfangreichen Infotainmentsysteme moderner Autos stellen dafür eine immer attraktivere Quelle für Ablenkungen aller Art dar.

Doch woher weiß man eigentlich, wie ablenkend oder wie gefährlich eine bestimmte Nebentätigkeit im Auto ist? Wie untersucht man wissenschaftlich, welches Ablenkungspotenzial von einem System ausgeht und wie sich dies auf die Fahrleistung auswirkt? An Studien herrscht jedenfalls kein Mangel. Sie stellen die Datengrundlage für politische und wirtschaftliche Entscheidungen dar.

Ein Beispiel für politische Entscheidungen: Als sich Ende der 90er Jahre Mobiltelefone rasant verbreiteten und auch während des Autofahrens genutzt wurden, gab es in der Folge zahlreiche Untersuchungen, die sich mit der Gefährlichkeit des Telefonierens am Steuer (mit oder ohne Freisprecheinrichtung) befassten. Als Konsequenz dieser Studien verbot der Gesetzgeber das Telefonieren (ohne Freisprecheinrichtung) im Jahr 2004 und verschärfte 2014 noch einmal die Regelung. Seither ist es noch nicht einmal mehr erlaubt, während des Fahrens einen eingehenden Anruf auf seinem Handy wegzudrücken. Die Begründung: Wissenschaftliche Studien haben gezeigt, dass dies zu sehr vom Fahren ablenkt.

Ein Beispiel für wirtschaftliche Entscheidungen: Für jedes neue Informations- oder Assistenzsystem im Auto muss der Automobilhersteller aufgrund einer Selbstverpflichtung der Automobilindustrie nachweisen, dass dieses System den Fahrer nicht zu stark vom Fahren ablenkt. Die genauen Regeln sind in den sogenannten *AAM-Guidelines* der *Alliance of Automobile Manufacturers* festgelegt. Der Vergleichsmaßstab ist hierbei die Bedienung eines konventionellen Autoradios (Alliance of Automobile Manufacturers, 2006, S. 46), und als gerade noch akzeptabel gelten beispielsweise Blickabwendungsdauern (weg von der Straße) von maximal zwei Sekunden. Man kann sich leicht ausmalen, wie groß der Druck ist, dass ein System, in das ein Hersteller mehrere Jahre Entwicklungszeit und hohe Kosten investiert hat, in einem solchen Test nicht durchfällt.

Dies zeigt: Die Ergebnisse solcher Untersuchungen zum Fahrverhalten haben auf unseren Lebensalltag erhebliche Auswirkungen. Sie beeinflussen Sicherheit, Komfort und Freiheit beim Autofahren sowie die Markteinführung neuer technischer Systeme im Fahrzeug mit den entsprechenden wirtschaftlichen Konsequenzen. Grund genug also, um sich die Frage zu stellen: Wie werden eigentlich solche Studien gemacht?

Mit welchen Methoden wird untersucht, ob ein System beispielsweise vom Fahren ablenkt?

Grob unterteilen lassen sich die Methoden zur Untersuchung des Fahrverhaltens in zwei Bereiche:

1. Studien in einem normalen Auto, das im realen Straßenverkehr fährt
2. Studien in einer simulierten Fahrumgebung

Innerhalb jedes Bereichs lässt sich noch weiter differenzieren, z. B. wie stark bei einer Fahrt im echten Straßenverkehr das Fahren kontrolliert wird, oder welchen Aufwand man bei der Simulation einer virtuellen Fahrumwelt im Labor treibt und wie realitätsnah dadurch das simulierte Fahren ist.

Die Unterschiede dieser verschiedenen Methoden der Fahrverhaltensbeobachtung sind gewaltig. Sie beeinflussen nicht nur die Durchführung einer Studie, sondern auch die damit untersuchbaren konkreten Fragestellungen oder die Art der erhobenen Daten – und damit selbstverständlich auch das Ergebnis der Studie. Am Ende klebt jedoch auf jedem neuen Assistenzsystem oder Anzeige- und Bedienkonzept für das Auto das Etikett „wissenschaftlich getestet“ – unabhängig davon, wie die Ergebnisse eigentlich zustande kamen. In Wahrheit sind jedoch die Ergebnisse in hohem Maß von diesen Methoden abhängig.

Im Folgenden werden daher die verschiedenen Methoden der Fahrverhaltensbeobachtung beschrieben und ihre unterschiedlichen Stärken und Schwächen aufgezeigt. Die Zusammenstellung verfolgt zwei Ziele: Zum einen soll sie dem interessierten Laien ein methodisches Hintergrundwissen vermitteln, um die Ergebnisse verkehrspsychologischer Untersuchungen kritisch hinterfragen zu können. Zum anderen soll damit dem Experten und Praktiker ein zusammenfassender Überblick über Vor- und Nachteile der verschiedenen Untersuchungsmethoden gegeben werden. Er ist eine nützliche Entscheidungshilfe, um für den jeweiligen Untersuchungszweck die jeweils geeignetste Methode auszuwählen.

Doch egal für welche Methode zur Fahrverhaltensbeobachtung man sich entscheidet – am Ende müssen die Ergebnisse immer denselben wissenschaftlichen Gütekriterien standhalten, nämlich *Objektivität*, *Reliabilität*, *interne Validität*, *Inhaltsvalidität*, *externe Validität* und *Replizierbarkeit*. Werden diese Qualitätsstandards nicht ausreichend berücksichtigt, ist eine Studie in ihrer Aussagekraft eingeschränkt – im Extremfall sogar völlig wertlos. In [Teil 1](#) des Whitepapers zu den Methoden der Fahrverhaltensbeobachtung werden daher die wichtigsten Qualitätsstandards einer wissenschaftlichen Untersuchung beschrieben und mit Beispielen aus der Praxis der Fahrverhaltensbeobachtung illustriert. Im vorliegenden Teil 2 werden die verschiedenen Methoden der Fahrverhaltensbeobachtung vorgestellt und ihre Vor- und Nachteile anhand der in [Teil 1](#) beschriebenen Gütekriterien erläutert.

2. METHODEN DER FAHRVERHALTENSBEOBSACHTUNG

NATURALISTIC DRIVING STUDIES

Das entscheidende Merkmal von *Naturalistic Driving Studies (NDS)* ist, dass auf diese Weise erhobene Fahrdaten unter möglichst natürlichen Bedingungen erhoben werden. Die Wahrscheinlichkeit ist deshalb hoch, dass Fahrer bei der Teilnahme an solchen Studien tatsächlich auch ihr „normales“ Verhalten zeigen, unverfälscht von irgendwelchen ungewohnten oder künstlichen Versuchsbedingungen, wie es sonst häufig bei wissenschaftlichen Untersuchungen der Fall ist. Typisch für *NDS* sind daher folgende Merkmale:

- Die Studienteilnehmer fahren ihr eigenes Auto, das mit der notwendigen Technik ausgestattet ist, oder ein entsprechendes Fahrzeug, das ihnen für einen längeren Zeitraum zur Verfügung gestellt wird.
- Es gibt *keinen* Versuchsleiter, der im Auto mitfährt und irgendwelche Instruktionen zur Fahrt gibt. Die Teilnehmer fahren völlig frei wann, wohin, wie lange und wie sie wollen.
- Die Versuchsfahrzeuge sind unauffällig mit der notwendigen Technik ausgestattet und zeichnen (über den CAN-Bus des Autos) möglichst viele Parameter des Fahrverhaltens auf.
- Der Erhebungszeitraum beträgt meist Wochen oder Monate.

Der Fahrer ist also bei der Datenerhebung in seiner Freiheit in keiner Weise eingeschränkt. Dadurch sowie durch die unauffällige Technik und die langen Beobachtungszeiträume blendet der Fahrer nach einiger Zeit aus, dass sein Verhalten (anonymisiert und auf abstrahiertem Niveau) analysiert wird. Nach einer Eingewöhnungsphase verhält er sich im Straßenverkehr so wie immer und nutzt das Fahrzeug auf eine Weise, wie er es sonst auch tut. *Naturalistic Driving Studies (NDS)* sind daher ein Paradebeispiel für eine *nicht-reaktive* Messmethode, d. h. für eine Methode, bei der das gesamte Drumherum der Mess-Situation das zu messende Verhalten möglichst wenig verfälscht. *NDS*-Daten sind somit sehr realitätsnah. Die *externe Validität* (d. h. die Generalisierbarkeit der Ergebnisse auf das alltägliche Fahrverhalten) solcher Untersuchungen ist deswegen sehr hoch (vgl. [Teil 1](#)).

Doch diese Unverfälschtheit und Natürlichkeit wird mit einem hohen wissenschaftlichen Preis erkaufte. Die Kehrseite der Medaille ist nämlich, dass dadurch auch jede Standardisierung und Kontrolle fehlt und somit die *interne Validität* gering ist. Möchte man beispielsweise wissen, wie sich ein Abstandsregeltempomat auf die gefahrenen Geschwindigkeiten oder den Sicherheitsabstand auswirkt, dann hat man beispielsweise folgende Möglichkeit der Analyse:

Man vergleicht für bestimmte Nutzungsfälle (= *use cases*, z. B. auf der Autobahn) alle Fahrsequenzen, in denen der Abstandsregeltempomat eingeschaltet war, mit den Sequenzen, in denen er *nicht* eingeschaltet war. Dazu berechnet man für diese beiden Bedingungen jeweils Durchschnitts- oder Maximalwerte von bestimmten Kennwerten (z. B. Geschwindigkeit, *Time-to-Collision*) und vergleicht sie statistisch miteinander. Der Haken dabei ist jedoch, dass man – mangels Kontrolle – keinen Einfluss darauf hat, in welchen Situationen der Fahrer sein System einschaltet oder ausschaltet. Gewiss tut er dies nicht per Zufall, sondern abhängig von bestimmten äußeren Umständen (Verkehrsaufkommen, Lichtverhältnissen, Sichtbedingungen) oder psychischen Umständen (Müdigkeit, Stress, Fahrspaß, Ablenkung durch eine Nebentätigkeit wie Essen am Steuer). Angenommen, ein Fahrer nutzt den Abstandsregeltempomat bevorzugt im dichten Berufsverkehr, nicht jedoch bei freier Strecke mit zahlreichen Überholmöglichkeiten: Wertet man die gefahrenen Durchschnittsgeschwindigkeiten aus, würde man

den falschen Schluss ziehen, dass ein Abstandsregeltempomat dazu führt, dass ein Fahrer damit deutlich langsamer fährt. In Wirklichkeit lagen die Geschwindigkeitsunterschiede jedoch nicht am Assistenzsystem, sondern am Verkehrsaufkommen – im dichten Berufsverkehr kann man eben nicht schnell fahren, ob mit oder ohne Assistenzsystem. Der gemessene Unterschied bei der gefahrenen Geschwindigkeit ist nicht eindeutig interpretierbar – ein Problem geringer *interner Validität*. Kausale Schlussfolgerungen sind daher bei *NDS* kaum möglich (vgl. [Teil 1](#)).

Dies macht das Hauptproblem von *NDS* deutlich: Die große Schwierigkeit liegt nicht etwa in der Aufzeichnung der Fahrdaten, sondern in ihrer anschließenden Analyse. Denn ohne Standardisierung und Kontrolle – das Wesen eines ordentlich durchgeführten Experiments – besteht die Herausforderung darin, solche Situationen aus der Flut an Fahrdaten zu identifizieren, die miteinander überhaupt vergleichbar sind. Ohne vergleichbare situative Bedingungen sind Fehlschlüsse bei den Auswirkungen von technischen Systemen auf das Fahrverhalten vorprogrammiert, denn man vergleicht dann Äpfel mit Birnen. Auch das Argument, dass sich solche Fehler am Ende bei Datenanalyse schon irgendwie „herausmitteln“ würden, zählt nur bedingt. Denn Fehler heben sich bei einer Berechnung von Durchschnittswerten nur dann gegenseitig auf, wenn sie *zufällig* verteilt sind. Liegt jedoch ein *systematischer* Fehler vor (= *bias*), also eine systematische Verzerrung der Daten, gilt dies keineswegs, z. B. das Einschalten des Abstandsregeltempomats nur im dichten Verkehr, wenn man ohnehin nicht schnell fahren kann.

Um also bei *NDS* zu sinnvollen Ergebnissen zu kommen, die nicht von Artefakten verfälscht sind, ist es besonders wichtig, Daten vor ihrer Analyse entsprechend zu filtern, beispielsweise beim Fahren auf einer Autobahn nur solche Situationen zu betrachten, die hinsichtlich bestimmter Kriterien gleich sein (z. B. gleiche Anzahl an Fahrspuren, gleiche Geschwindigkeitsbeschränkung, gleiches Verkehrsaufkommen, gleiche Witterungsverhältnisse usw.). Je nachdem welche Filterkriterien man anwendet, werden jedoch auch die Ergebnisse anders ausfallen. Da es keine festgesetzten Regeln oder Konventionen gibt, welche Filter anzuwenden sind, liegen diese Entscheidungen im Ermessen desjenigen, der die Daten auswertet. Dies ist nicht unproblematisch, denn es verstößt gegen ein grundlegendes Qualitätsmerkmal wissenschaftlichen Arbeitens, nämlich gegen die *Objektivität*. Wenn Studienergebnisse davon abhängen, welche Person die Studie auswertet, dann ist die Studie nicht objektiv und damit auch nur begrenzt aussagekräftig (vgl. [Teil 1](#)).

Dass *Naturalistic Driving Studies* ein Objektivitätsproblem haben, mag für viele überraschend klingen, denn das Image von *NDS* suggeriert das glatte Gegenteil. Allgemein lässt sich wohl sagen: Je mehr Technik bei der Aufnahme von Verhaltensdaten zum Einsatz kommt, desto objektiver und präziser ist das Image der Methode. Bei *NDS* wird zudem noch nicht einmal das menschliche Verhalten selbst gemessen, sondern das „Verhalten“ einer Maschine (= Auto) protokolliert, die von einem Menschen bedient wird. Doch darf man eben nicht aus dem Auge verlieren, dass sich *NDS*-Daten nicht von selbst auswerten. Am Ende sitzen stets Menschen vor einer Datenbank mit vielen Terabyte an Fahrdaten. Sie formulieren Hypothesen, wählen *use cases* aus und definieren Kriterien zur Filterung und Zusammenfassung (= Aggregation) von Daten. Und hierbei gibt es sehr viele Freiheitsgrade und Ermessensspielräume.

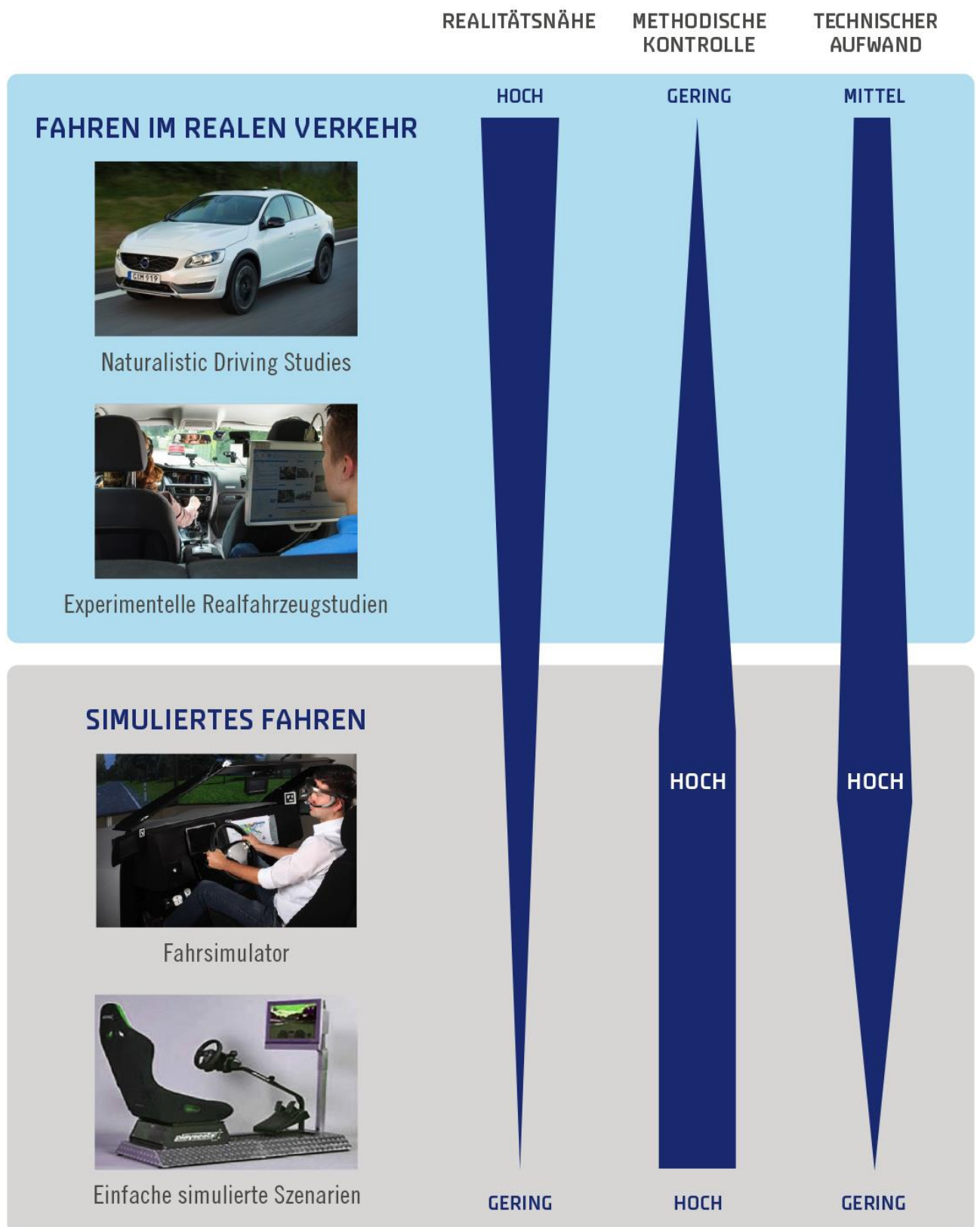


Abbildung 1: Realitätsnähe des Fahrens, methodische Kontrolle und technischer Aufwand der verschiedenen Arten der Fahrverhaltensbeobachtung im Überblick.

Zudem ist die statistische Analyse solch enormer Datenmengen mit einem starken „Daten-Rauschen“ enorm kompliziert. Es braucht hierfür ein interdisziplinäres Team aus Informatikern mit Spezialisierung auf Data-Mining-Methoden, Statistikern mit Spezialkenntnissen über anspruchsvolle statistische Analysemethoden großer Datenmengen und Ingenieuren (da die gemessenen Fahrdaten auch korrekt interpretiert werden müssen).

Dennoch sind *Naturalistic Driving Studies* eine wichtige und wertvolle Methode, wenn man sie für die richtigen Fragestellungen einsetzt. Weniger geeignet sind sie für eine Untersuchung der Auswirkungen von Assistenzsystemen auf das Fahrverhalten. Ihre Stärke sind vielmehr explorative Fragestellungen: Möchte man beispielsweise wissen, welche Systeme oder Systemfunktionen eines Autos ein Fahrer in welchen Situationen von sich aus freiwillig nutzt, dann lässt sich dies nur mit einer *NDS* sinnvoll untersuchen. Denn der Fahrer braucht hierfür die Freiheit, selbst entscheiden zu können, welches System er wann auf welche Weise nutzt. In einer experimentellen Studie im Realfahrzeug oder Fahr Simulator hingegen wird ihm aufgrund der beabsichtigten Standardisierung und Kontrolle über die Versuchsbedingungen viel zu viel vorgegeben, so dass dadurch individuelle Vorlieben und Gewohnheiten sowie auf die Verkehrssituation angepasstes Nutzungsverhalten nicht sinnvoll untersuchbar ist.

EXPERIMENTELLE REALFAHRZEUGSTUDIEN

Experimentelle Realfahrzeugstudien (= experimentelle Feldstudien, *experimental Field Operational Tests*, *eFOT*) finden im Gegensatz zu *Naturalistic Driving Studies* (*NDS*) unter stärker kontrollierten Bedingungen statt. „Kontrollierte Bedingungen im Straßenverkehr“ sind allerdings für Forscher, die es gewohnt sind, Fahrexperimente im Simulator durchzuführen, schon ein Widerspruch in sich. Denn ganz wesentliche Einflussgrößen sind im realen Straßenverkehr eben *nicht* kontrollierbar, v. a. das Verhalten anderer Verkehrsteilnehmer. Entsprechend versucht man, das Beste aus den gegebenen Bedingungen zu machen und den Ablauf der Untersuchung zu standardisieren, wo es möglich ist, beispielsweise:

- Es wird eine Fahrtstrecke vorgegeben.
- Die zu fahrende Route wird ins Navigationssystem einprogrammiert und von diesem angesagt (anstatt vom Versuchsleiter), um bei den Navigationshinweisen stets gleiche Bedingungen bei allen Probanden zu haben.
- Es gibt verschiedene experimentelle Bedingungen. Wenn z. B. die Auswirkungen eines Fahrerassistenzsystems auf das Fahrverhalten getestet werden sollen, durchfährt jeder Proband dieselbe Strecke einmal mit System (= Versuchsbedingung) und einmal ohne (= Kontrollbedingung) bzw. er durchfährt eine möglichst vergleichbare Strecke, um Lerneffekten vorzubeugen.

Doch trotz noch so großer Anstrengungen wird man im realen Straßenverkehr niemals identische Bedingungen schaffen können. Durchfahren 50 Probanden dieselbe Stelle (z. B. eine große Kreuzung), wird in Abhängigkeit vom Verhalten anderer Verkehrsteilnehmer die Verkehrssituation immer ein wenig anders sein.

Ein weiteres simples und pragmatisches Unterscheidungskriterium zwischen einer *Naturalistic Driving Study* (*NDS*) und einer experimentellen Realfahrzeugstudie (*eFOT*) ist der Versuchsleiter. Bei einer *NDS* ist grundsätzlich kein Versuchsleiter anwesend und notwendig – der Fahrer setzt sich einfach zu einem beliebigen Zeitpunkt ins Auto und fährt, wohin er will, während das Fahrzeug automatisch die Fahrdaten speichert. Bei einer Realfahrzeugstudie hingegen sitzt der Versuchsleiter mit im Auto, bedient die

Technik, gibt Instruktionen, führt evtl. Verhaltensbeobachtungen mit Beobachtungsbögen durch oder stellt Fragen (z. B. in kritischen Situationen oder bei Eingreifen eines Assistenzsystems).

Die experimentellen Realfahrzeugstudien haben gegenüber den *NDS* den Vorteil, dass durch die Anwesenheit eines Versuchsleiters auch Methoden eingesetzt werden können, die weit mehr über die kognitiven Prozesse des Fahrers verraten, als dies durch das bloße Aufzeichnen von Fahrdaten bei *NDS* möglich ist, beispielsweise:

- Eye-Tracking, um präzise messen zu können,
 - *wo genau* der Fahrer hinsieht (Verkehrsgeschehen, Fahrerinformationssysteme, Bedieneinrichtungen)
 - zu *welchem Zeitpunkt* er wohin sieht (Bemerken von Gefahrensituationen)
 - *wie lange* er wohin sieht (Ablenkungswirkung von Informationssystemen)
- Physiologische Messungen (z. B. Herzrate, elektrische Hautleitfähigkeit, EEG), um Rückschlüsse über die mentale Beanspruchung des Fahrers in verschiedenen Verkehrssituationen oder Versuchsbedingungen ziehen zu können
- Fragebögen zu bestimmten Verkehrssituationen, die zwar gern als *paper&pencil*-Methode belächelt werden, sich aber in der Praxis bei der Datenauswertung oft als aussagekräftiger erweisen, als manch mit großem Hightech-Aufwand erhobener Parameter

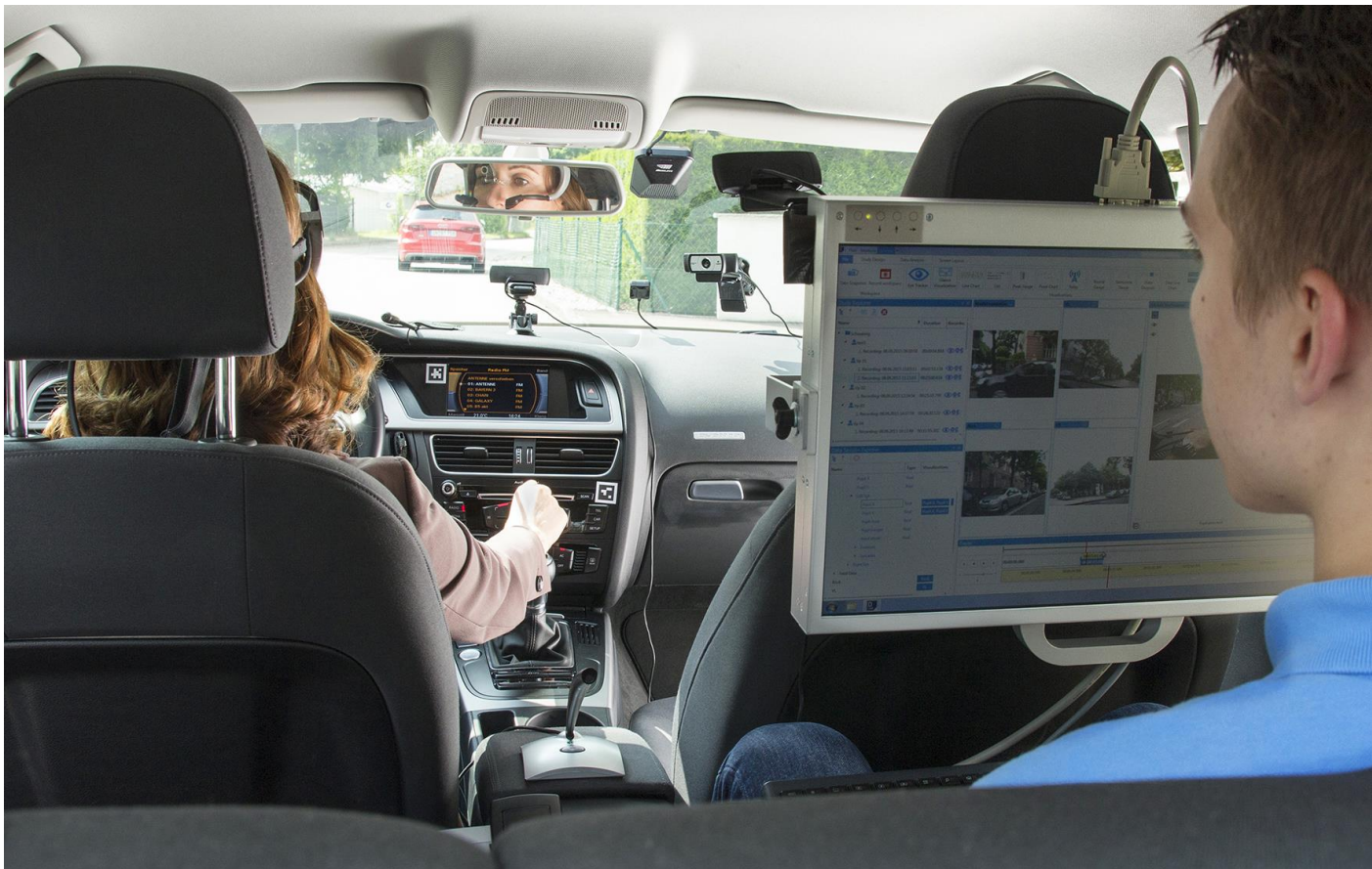


Abbildung 2: Im Gegensatz zu *Naturalistic Driving Studies* fährt bei experimentellen Realfahrzeugstudien immer ein Versuchsleiter mit, der die Technik überwacht, Instruktionen gibt und gegebenenfalls Fahrfehler protokolliert und Befragungen während der Fahrt durchführt.

Der technische Aufwand für eine Realfahrzeugstudie ist sehr hoch, vor allem wenn nicht nur (wie bei *NDS*) Fahrdaten über den CAN-Bus-Anschluss aufgezeichnet werden sollen, sondern gleichzeitig auch Videoströme und Daten von Blickverhalten oder physiologischen Messungen. Schnell führt dies zu einem Kofferraum voll Messtechnik, die nur von Spezialisten eingebaut und bedient werden kann, auch wenn es mittlerweile einfach zu handhabende Komplettlösungen gibt.



Abbildung 3: Früher waren Verhaltensstudien im realen Verkehr stets verbunden mit einem Kofferraum voll Messtechnik (Bild links). Mittlerweile gibt es jedoch einfach zu handhabende Komplettlösungen, die auf dem Rücksitz Platz haben (Bild rechts).

Die entscheidenden Vorteile sind hingegen, dass das Fahren in einem solchen Versuchsfahrzeug zwar nicht ganz das natürliche Setting bei einer *Naturalistic Driving Study* erreicht, aber dennoch um Längen realitätsnäher ist als das Fahren in einem Fahrsimulator. Der Grad an experimenteller Kontrolle ist immerhin mittel – nicht so hoch standardisiert wie in einem Fahrsimulator, aber eben auch nicht völlig unkontrolliert wie bei einer *NDS*. Zusätzlich ermöglichen weitere Methoden wie Eye-Tracking oder Fragebögen, nicht nur zu beschreiben, *wie* sich ein Fahrer verhält, sondern auch zu verstehen, *warum* er sich so verhält.

Ein Nachteil einer experimentellen Realfahrzeugstudie ist, dass durch die sehr strikten und detaillierten Vorgaben eines mitfahrenden Versuchsleiters und durch die deutlich sichtbare Messtechnik den Probanden in hohem Maße bewusst ist, dass sie gerade an einer Studie teilnehmen und ihr Verhalten beobachtet wird. Deswegen neigen sie eher dazu, sozial erwünschtes Verhalten zu zeigen. Das heißt, sie fahren tendenziell langsamer, rücksichtsvoller und regelkonformer – ähnlich wie bei einer Fahrprüfung – obwohl sie im Alltag möglicherweise ganz anders fahren. Die *externe Validität* ist daher zwar insofern gegeben, dass die Probanden ein echtes Auto im realen Straßenverkehr fahren (im Gegensatz zu den Bedingungen im Fahrsimulator), aber andererseits dadurch eingeschränkt, dass sich das beobachtete Fahrverhalten (im Gegensatz zu *NDS*), nicht so ohne weiteres auf das alltägliche, normale Fahrverhalten übertragen lässt.

EXPERIMENTE IM FAHRSIMULATOR

Der größte Vorteil im Fahrsimulator ist das hohe Ausmaß an Kontrolle, die der Versuchsleiter über den gesamten Versuchsablauf und die erhobenen Daten hat. Nahezu alle fahrrelevanten Einflussgrößen können im Simulator kontrolliert und manipuliert werden: Die zu fahrende Strecke, der Fremdverkehr (einschließlich Geschwindigkeiten anderer Fahrzeuge), die technische Ausstattung des Fahrzeugs mit Assistenzsystemen unterschiedlicher Auslegung sowie Beleuchtungs- und Wetterverhältnisse. Anders als beim Fahren im realen Straßenverkehr ist es daher möglich, methodisch saubere experimentelle Untersuchungen durchzuführen, bei denen gezielt nur eine oder mehrere Einflussgrößen (= unabhängige Variablen) verändert werden, während alle sonstigen Bedingungen konstant bleiben. Dadurch ist es möglich, eindeutige kausale Schlussfolgerungen zu ziehen und die gemessenen Unterschiede beim Fahrverhalten eindeutig kausal auf veränderte Bedingungen (z. B. Fahren mit oder ohne Assistenzsystem) zurückzuführen. All dies wirkt sich positiv auf die *interne Validität* aus (vgl. [Teil 1](#)).



Abbildung 4: Bei einem hochwertigen Fahrsimulator (hier vom Deutschen Zentrum für Luft- und Raumfahrt, DLR) sitzt der Fahrer in einem echten Auto. Die virtuelle Fahrumwelt wird dabei rund um das Auto auf eine Leinwand projiziert. Der gesamte Aufbau ist auf sechs beweglichen Hydraulikstelzen (= Hexapod) gelagert, die durch Bewegung der gesamten Kabine typische beim Autofahren auftretende Vibrationen, Beschleunigungen und Fliehkräfte imitieren sollen. Bei manchen Versuchspersonen führt dies zu Übelkeit.

Ein weiterer großer Vorteil ist die hohe Datenqualität: Weil die Fahrumgebung vollständig im Computer erzeugt wird, sind automatisch auch alle Umgebungsfaktoren (z. B. Abstände zu anderen Fahrzeugen oder Spurpositionen) bekannt und müssen nicht erst durch (fehleranfällige) Sensoren gemessen werden. Auch das gesamte Timing ist kontrollierbar und automatisch erfassbar. Man weiß beispielsweise, zu welchem Zeitpunkt ein Signal erscheint (z. B. eine Ampel auf Rot schaltet), man kann einen Fußgänger genau dann die Straße überqueren lassen, wenn der Fahrer eine bestimmte (unsichtbare) Streckenmarkierung überfährt, oder kann das vorausfahrende Fahrzeug (mit einer genau vorgegebenen Verzögerung) bremsen lassen, wenn ein definierter Sicherheitsabstand (*Time to Collision*) vorhanden ist.

Sowohl die hohe experimentelle Kontrolle als auch die hohe Datenqualität bei der Erfassung der Reizdarbietung (= was der Fahrer zu welchem Zeitpunkt zu sehen bekommt) und des Fahrverhaltens (= wie der Fahrer auf die dargebotenen Reize reagiert) führen zu einer vergleichsweise hohen Messgenauigkeit (= *Reliabilität*). Dadurch ist auch die *interne Validität* relativ hoch, d. h. man kann davon ausgehen, dass ein solches Experiment im Simulator auch das misst, was man eigentlich messen möchte (z. B. die Auswirkungen eines Assistenzsystems) und nicht etwa die Auswirkungen anderer Einflussgrößen des Verkehrs (vgl. [Teil 1](#)).

Es lassen sich im Simulator selbstverständlich auch zusätzliche Methoden einsetzen, die über das reine Fahrverhalten hinausgehen und mit denen man tieferen Einblick in die Wahrnehmung und mentale Beanspruchung des Fahrers bekommt, z. B. Eye-Tracking und physiologische Messungen. Unter solchen Laborbedingungen ist die Qualität dieser Daten sogar noch höher als bei einer Fahrstudie im realen Straßenverkehr. Beispielsweise gibt es im Fahrsimulator kein störendes Sonnenlicht, das durch seinen hohen Infrarot-Anteil die Pupillenerkennung von Eye-Tracking-Systemen stört und dadurch die Qualität von Blickdaten beeinträchtigt. Unter den standardisierten Lichtverhältnissen im Labor liefern Blickerfassungsbrillen die qualitativ besten Ergebnisse.

Der Hauptkritikpunkt an Fahrsimulator-Studien ist jedoch ihre begrenzte Realitätsnähe. Selbstverständlich gibt es erhebliche Unterschiede zwischen dem Fahren in einem Simulator und dem Fahren in einem echten Auto im realen Straßenverkehr. Zudem gibt es gewaltige technische Unterschiede zwischen Fahrsimulatoren. Sie reichen von extrem aufwändigen, dynamischen Simulatoren mit 360-Grad-Sichtsimulation, einer originalgetreuen Fahrzeugkabine und einer sehr realistischen grafischen Darstellung der virtuellen Fahrumwelt bis hin zu ganz einfachen, statischen Sitzkisten mit einem simplen Monitor für die visuelle Darbietung der Fahrumwelt.

Die eingesetzte Technik hat großen Einfluss auf das Fahrverhalten und damit auf die Ergebnisse einer Fahrstudie. Möchte man beispielsweise den Einfluss von Fahrerassistenzsystemen testen, die in die Längsführung des Fahrzeugs eingreifen, also bremsen oder beschleunigen (z. B. Abstandsregeltempomat oder Notbremsassistent), macht es wenig Sinn, dies in einem statischen Fahrsimulator zu untersuchen. Denn anders als in einem richtigen, fahrenden Auto spürt der Fahrer dort nicht die wirkenden Kräfte auf seinen Körper, die ihn beim Beschleunigen nach hinten in den Sitz und beim Bremsen nach vorn in den Gurt drücken. Dem Fahrer fehlt dadurch das unmittelbare, natürliche Feedback. Dass sein Fahrzeug gerade automatisch eine starke Bremsung durchführt, ist für ihn nur visuell über eine Verlangsamung des optischen Flusses auf den Monitoren wahrnehmbar – und das funktioniert sehr schlecht.

Fehlt dem Fahrer im Simulator jedoch ein realistisches Fahrgefühl, dann fährt er auch ganz anders, als er dies in Realität tun würde. Selbst bei einem dynamischen Fahrsimulator hat die Art und Weise, wie man Fahrzeugbeschleunigungen simuliert, einen großen Einfluss auf die gemessenen Fahrverhaltensdaten. Das zeigt beispielsweise die

Untersuchung von Dagelen et al. (2009), in der verschiedene Techniken und Algorithmen zur Simulation von Fahrdynamik miteinander verglichen wurden.

Natürlich wäre es wünschenswert, für Untersuchungen nur Fahrsimulatoren zu verwenden, die ein möglichst realistisches Fahrgefühl vermitteln. Doch weil der technische Aufwand und die damit verbunden Kosten immens hoch sind, überwiegen die Studien, die mit einfacheren, weniger realistischen Fahrsimulatoren gemacht wurden.



Abbildung 5: Ein einfacher, statischer Fahrsimulator, bei dem die Fahrumgebung auf drei großen Monitoren dargestellt wird.

In Fahrsimulatoren tritt noch ein weiteres Problem auf: Die Simulatorkrankheit (Straus, 2005; Regan & Price, 1994). Manchen Teilnehmern wird vom Fahren in einem Simulator schlecht. Die Übelkeit entsteht, weil das Gehirn einander widersprechende Informationen verarbeiten muss (Guedry, 1970). Beispielsweise melden die Augen beim Durchfahren einer Kurve an das Gehirn, dass man sich bewegt, das Gleichgewichtsorgan meldet jedoch keine Bewegung, da keine Fliehkräfte wirken. Das Problem liegt also darin, dass die Simulation nicht perfekt ist, sondern manche Aspekte realistisch simuliert sind, andere jedoch schlecht oder gar nicht. Paradoxe Weise tritt jedoch die Simulatorkrankheit häufiger bei guten und aufwändigen Fahrsimulatoren mit Bewegungssimulation auf als bei einfachen, statischen, vermutlich deswegen, weil der Fahrer dadurch tiefer in die simulierte Welt eintaucht (= höherer Immersionsgrad), das Gehirn jedoch widersprüchliche Rückmeldungen (z. B. nicht völlig stimmige Beschleunigungssimulationen) trotzdem bemerkt und – im wahrsten Sinne – übelnimmt. Je nach zu fahrender Strecke im Simulator (Abbiegesituationen in simulierten Ortschaften sind beispielsweise

unangenehmer als Autofahrten) gibt es daher einen teils erheblichen Prozentsatz an Probanden, die den Fahrversuch wegen Übelkeit abbrechen müssen.

Von gelegentlicher Übelkeit abgesehen ist jedoch das Fahren in einem Fahrsimulator ohne jedes Risiko. Dies ist Vor- und Nachteil zugleich. Einerseits lassen sich dadurch im Simulator viele Dinge untersuchen, die im realen Straßenverkehr viel zu gefährlich wären, da das Unfallrisiko juristisch und ethisch nicht zu verantworten wäre. Andererseits führt jedoch die Ungefährlichkeit der Untersuchungssituation dazu, dass Versuchspersonen im Fahrsimulator Verhaltensweisen an den Tag legen, die ihnen im realen Straßenverkehr viel zu riskant wären. Gerade für eine häufige Fragestellung im Fahrsimulator, der Evaluation von Anzeige- und Bedienkonzepten im Fahrzeug auf ihre Ablenkungswirkung, kann man daher bei Simulator-Experimenten außergewöhnlich lange Blickabwendungszeiten von der Straße beobachten oder dramatische Fahrfehler wie ein Abkommen auf die Gegenfahrbahn.

Was man im realen Verkehr schnell mit dem Leben bezahlt, hat im Simulator keine negativen Konsequenzen. Weil die Probanden das wissen, widmen sie deshalb mehr Aufmerksamkeit auf die Bearbeitung einer Sekundäraufgabe (z. B. Bedienung eines Assistenzsystems) zulasten der Primäraufgabe (das Fahren selbst, z. B. Spurhaltung und Abstandsregelung). Bei der Datenanalyse erkennt man dies dann beispielsweise einerseits an kürzeren Gesamtbedienzeiten des getesteten Systems und längeren Blicken auf ein Display, die seltener von Kontrollblicken auf die Straße unterbrochen werden, als dies in der Realität der Fall wäre. Konsequenterweise kommt es dabei auch zu mehr seltenen, kritischen Verkehrssituationen (Beinahe-Kollisionen und Unfälle) als im echten Straßenverkehr. Deswegen ist es problematisch, von Experimenten im Fahrsimulator auf das Verhalten von Fahrern im realen Straßenverkehr zu schlussfolgern, denn die gesamte Untersuchungssituation ist hochgradig *reaktiv*, d. h. die Messsituation beeinflusst das normale Verhalten der Probanden stark. Das Verhalten im Fahrsimulator ist nicht repräsentativ für das Verhalten im realen Straßenverkehr, oder, wie der Wissenschaftler sagt: Im Simulator gewonnen Ergebnissen mangelt es an *externer Validität* (vgl. [Teil 1](#)).

Das Defizit an *externer Validität* wird zusätzlich durch dieselben Faktoren verstärkt, die auch schon für praxisnahe Realfahrzeugstudien genannt wurden: Auch im Fahrsimulator ist den Probanden natürlich bewusst, dass ihr Verhalten beobachtet und gemessen wird. Sie verhalten sich daher verkehrsregelkonformer und sozial erwünschter als sie dies möglicherweise im Alltag tun. Das natürliche Fahrverhalten kann man daher besser mit *Naturalistic Driving Studies* untersuchen.

Obwohl die geringe *externe Validität* von Fahrsimulator-Studien seit Langem beklagt wird, gibt es dennoch nur wenige Untersuchungen, die systematisch die Abweichungen von Simulator-Daten von im Realverkehr gewonnenen Daten erforschten. Eine Ausnahme ist die Studie von Engström, Johansson und Östlund (2005). Sie zeigt beispielsweise, dass der physiologische Workload von Fahrern und ihre Lenkaktivität im Realfahrzeug höher sind, als dies in Simulator-Versuchen der Fall ist. Die Autoren begründeten dies mit dem höheren objektiven Risiko im realen Straßenverkehr. In einer anderen Studie (Victor, Harbluk & Engström, 2005) wurde untersucht, wie sich die Effekte von visuellen und auditiven Anforderungen während der Fahrt auf das Blickverhalten von Fahrern auswirken. Dazu wurden die Experimente in zwei statischen Simulatoren und in einem Realfahrzeug durchgeführt und miteinander verglichen. Das Blickverhalten in den beiden Simulator-Bedingungen war dabei ähnlich, unterschied sich jedoch deutlich vom Blickverhalten im echten Verkehr. Dies lässt den Schluss zu, dass im Fahrsimulator erhobene Blickdaten *extern nicht valide* sind. Denn offensichtlich zeigen Fahrer im Simulator ein anderes Blickverhalten als beim Autofahren im realen Verkehr (Lietz et al., 2011).

Zusammenfassend bleibt festzuhalten, dass Datenerhebungen in einem Fahrsimulator durch die hohe experimentelle Kontrollierbarkeit, die Standardisierung und die präzise Messung von Daten zwar eine hohe *interne Validität* besitzen. Die Interpretation der Ergebnisse ist jedoch schwieriger, als es auf den ersten Blick scheint, und die *externe Validität* ist gering, weil im Simulator gewonnene Ergebnisse nicht ohne weiteres auf den realen Straßenverkehr übertragbar sind (vgl. [Teil 1](#)).

EINFACHE FORMEN SIMULIERTEN FAHRENS

Darüber, was ein Auto ist und was nicht, gibt es einen Konsens – obwohl die Vielfalt an Autos sehr groß ist. Und welche Eigenschaften ein Auto haben muss, um eine Zulassung zum Straßenverkehr zu bekommen, ist in Deutschland in der Fahrzeug-Zulassungsverordnung (FZV) gesetzlich geregelt. Ab wann man jedoch eine Anordnung von Technik-Komponenten als Fahrsimulator bezeichnen kann, darüber gehen die Meinungen auseinander. Und welche Eigenschaften ein solcher Simulator haben muss, um damit sinnvoll Untersuchungen zum Fahrverhalten durchzuführen, ist nirgendwo definiert. Was daher für die einen ein Fahrsimulator ist, ist für andere bloß eine „Sitzkiste“ oder ein *Mock-up*. Entsprechend werden auch nicht wenige Untersuchungen durchgeführt, die als Hardware für die Fahrsimulation lediglich einen PC-Monitor mit Spiele-Lenkrad und Spiele-Pedalerie aus dem Elektromarkt verwenden. Solche kostengünstigen Untersuchungen werden zwar meist nicht publiziert und auch überwiegend für Tests von *Entwürfen* von neuen Anzeige-/Bedienkonzepten durchgeführt, aber dennoch werden auf diese Art gewonnene Ergebnisse als Datengrundlage für Entscheidungen herangezogen – sonst wäre eine solche Untersuchung schließlich auch überflüssig.

Studien, die mit solchen „Primitiv-Fahrsimulatoren“ durchgeführt werden, haben jedoch einen langfristigen Nebeneffekt auf die gesamte Szene an Forschern und Entwicklern, die sich mit der Untersuchung des Fahrverhaltens befassen: Sie erweitern die Bandbreite an Studien, die nicht mehr viel mit dem echten Fahren im echten Straßenverkehr zu tun haben, noch einmal deutlich nach unten. Wenn schon Fahrstudien in einem sehr guten, dynamischen Fahrsimulator mit aufwändiger Sichtsimulation in ihrer *externen Validität* kritisch zu betrachten sind, dann gilt dies für Studien aus „Primitiv-Fahrsimulatoren“ natürlich umso mehr. Durch die Menge solcher Studien verschiebt sich mit der Zeit der Qualitätsmaßstab für Verhaltensbeobachtungen in der Verkehrspsychologie. Im Vergleich zu Fahrverhaltensmessungen in „Sitzkisten“ erscheint ein Experiment in einem mittelmäßigen, statischen Fahrsimulator auf einmal als methodisch vorbildlich und aussagekräftig. Doch der Maßstab, an dem sich eine Fahrstudie messen lassen muss, ist das Autofahren im realen Straßenverkehr und nicht die Bedienung eines Spielelenkrads vor einem PC-Monitor. Die bloße Existenz von Low-Tech-Fahrsimulationen reduziert bei Verkehrswissenschaftlern den Druck, den „normalen“ Fahrsimulator als Untersuchungsmethode zu rechtfertigen und die Generalisierbarkeit solcher Ergebnisse auf den Alltag kritisch zu hinterfragen. Denn man kann dadurch immer auf Untersuchungsbedingungen verweisen, die noch viel schlechter sind, d. h. noch unrealistischer, künstlicher und alltagsferner.

Bei Fahrsimulatoren kommt es jedoch nicht nur auf die verwendete Hardware an, sondern auch auf die Software, d. h. auf die Art der Simulation, die Realitätsnähe der virtuellen Objekte (Straßen, Gebäude, Fahrzeuge) und die konkret zu fahrenden Strecken und Verkehrsszenarien.

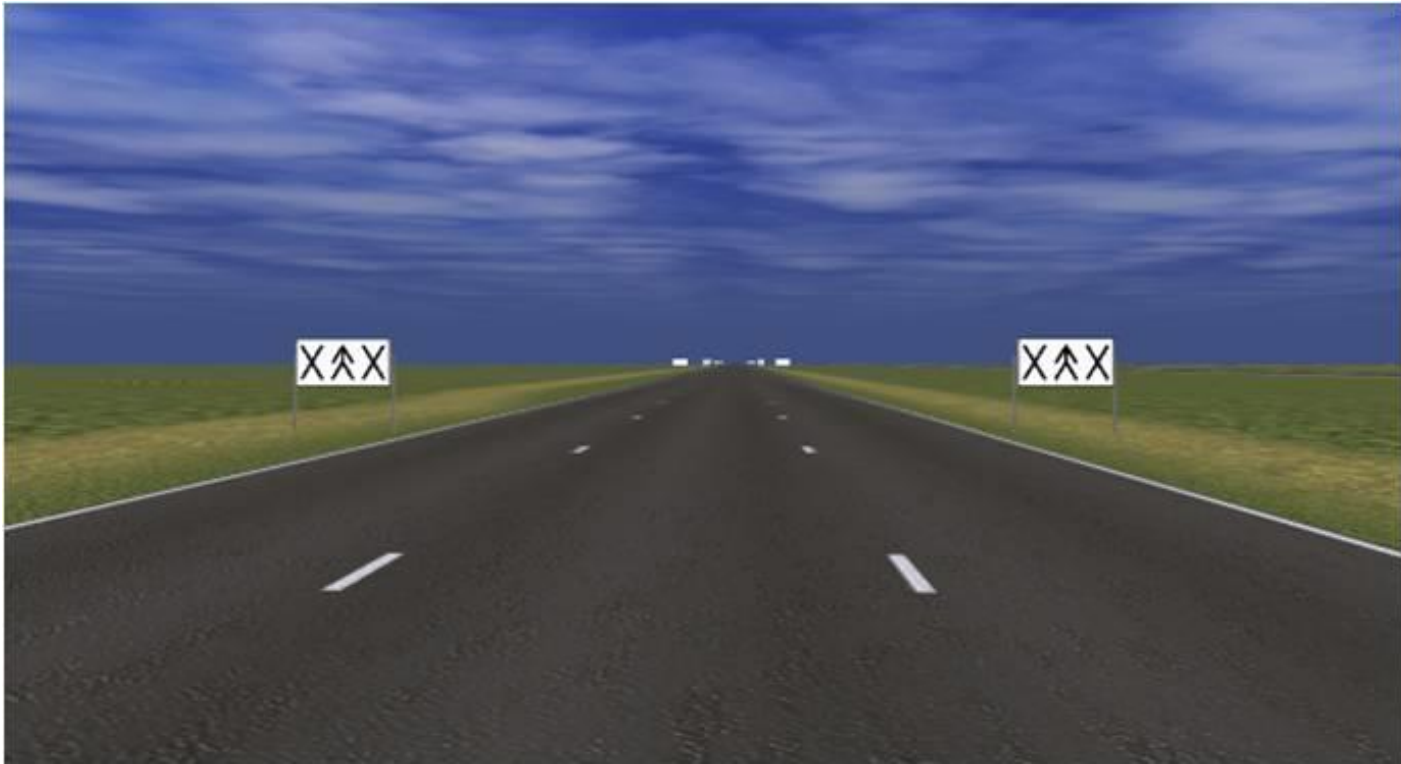


Abbildung 6: Fahrbahn, Wiese, Himmel. Es geht immer noch ein bisschen simpler: Screenshot aus dem Lane Change Task. Die in der Ferne auftauchenden Schilder geben dem Fahrer vor, auf welcher Fahrspur er fahren soll – hier auf der mittleren.

Ein besonders häufig verwendetes Szenario ist beispielsweise der sogenannte *Lane Change Task (LCT)*. Es handelt sich dabei um eine hoch standardisierte Fahraufgabe, die auf das Allernotwendigste reduziert ist: Eine dreispurige Fahrbahn und eine Vorgabe in Form von auftauchenden Verkehrsschildern, auf welcher der drei Fahrspuren der Fahrer fahren soll. Experimente werden hierbei nach dem Zweitaufgabenparadigma durchgeführt, d. h. dass zu dieser standardisierten Hauptaufgabe (= Spurwechsel beim Fahren) immer eine zu untersuchende Nebenaufgabe gegeben ist, z. B. die Bedienung eines Infotainmentsystems. Der Einfluss dieser Nebenaufgabe auf die Leistung in der Hauptaufgabe kann beim *LCT* analysiert werden, indem die Abweichung der beobachteten Fahrspur von einer idealen Fahrlinie gemessen wird. Der *LCT* wird vor allem in sehr frühen Produktentwicklungsstadien eingesetzt und besitzt einige große Vorteile: Er ist mit einem Spiele-Lenkrad an jedem PC durchführbar, er ist mit einem Minimum an Zeit und Kosten einsetzbar, er ist hoch standardisiert, leicht auszuwerten und liefert *reliable* und *reproduzierbare* Ergebnisse. Deswegen ist er bei vielen Akteuren recht beliebt: Automobilfirmen legen beispielsweise Wert auf geringe Kosten, Ingenieure auf einfach auszuwertende Tests und Psychologen auf experimentelle Kontrolle sowie *interne Validität*. Der *LCT* ist derart populär und weitverbreitet, dass es mittlerweile sogar eine eigene ISO-Norm gibt (ISO 26022, 2010), die das gesamte Verfahren (Aufbau, Durchführung und Auswertung solcher Studien) detailliert regelt.

Der große Haken am *Lane Change Task* ist jedoch seine schlechte *externe Validität*. Er hat mit Autofahren im realen Straßenverkehr oder gar im komplexen Verkehr einer Innenstadt nicht mehr viel zu tun. Er ist nur noch ein Computerspiel, bei dem es auf Reaktionsschnelligkeit und Geschicklichkeit ankommt. Entsprechend kann auch jedes Kind einen *LCT* fahren. Für den Erwerb eines Führerscheins braucht es jedoch nicht umsonst wochenlanges Training und für ein wirkliches sicheres Fahren Zehntausende Kilometer Fahrpraxis.

Der *Lane Change Task* ist damit ein recht anschauliches Beispiel für einen allgemeinen Trend in den Verhaltenswissenschaften: Ein äußerst komplexes Phänomen – hier das Autofahren im Straßenverkehr – wird durch permanente Abstrahierung auf einige ganz grundlegende Verhaltensmerkmale reduziert. Ganz wesentliche Merkmale, die einen Einfluss auf das zu messende Verhalten haben, werden für die Untersuchung des Phänomens einfach als Störfaktoren definiert und entweder während des Experiments konstant gehalten oder komplett aus der Untersuchungsmethode eliminiert (beim *LCT* z. B. der gesamte Fremdverkehr). Die komplexe Realität wird dadurch auf so wenige Einzelaspekte reduziert, dass man mit der gesamten Methodik nicht mehr das komplexe Phänomen selbst untersucht, sondern nur noch einzelne Teilaspekte davon – diese dafür jedoch umso tiefergehend und präziser. Abstrakt ausgedrückt: Man fixiert einseitig auf die *interne Validität* und verliert die *externe Validität* aus den Augen. Das, was der *Lane Change Task* misst (v. a. Reaktionsgeschwindigkeit und Feinmotorik), misst er sehr präzise, *reliabel* und *reproduzierbar* – aber es hat nicht mehr viel mit dem realen Autofahren zu tun.



Abbildung 7: Die reduzierteste Form der Fahrsimulation: Ein simpler Tischaufbau mit Spielelenkrad und Spielepedalerie für die „Fahrzeugsteuerung“ sowie PC-Monitor für die Sichtsimulation. Auf dem zweiten Monitor rechts daneben wird die Nebenaufgabe dargestellt. Für Untersuchungen zum *Lane Change Task* ist dies eine typische Versuchsanordnung.

Der *Lane Change Task* sei hier nur als besonders bekanntes Beispiel angeführt. Die Kritik hinsichtlich seiner Realitätsferne gilt jedoch ganz allgemein auch für ähnliche reduktionistische Methoden des simulierten Fahrens. Eine andere Methode ist beispielsweise die Instruktion an die Probanden, einem bestimmten vorausfahrenden Fahrzeug mit einem konstanten Abstand hinterher zu fahren. Als Zweitaufgabe ist dann beispielsweise ein Informationssystem zu bedienen. Gelegentlich bremst dann das Vorderfahrzeug überraschend. Gemessen wird dann die Schnelligkeit der Bremsreaktion und Konstanz des eingehaltenen Sicherheitsabstands als Maß für die Ablenkungswirkung der Zweitaufgabe.

Noch reduktionistischer ist die bloße visuelle Darbietung von Verkehrsszenarien als bewegte oder statische Bilder. Gemessen wird dann, wie viele Informationen der Proband von dieser Szenerie gesehen hat. Bei der Untersuchung von Hirnaktivitäten mit bildgebenden Verfahren (= funktionelle Magnetresonanztomographie) ist dies derzeit die einzig mögliche Methode der Reizdarbietung, denn die Versuchsperson liegt dabei ja in einem fMRT-Scanner. Die Frage ist trotzdem, ob man dabei etwas über Hirnaktivitäten beim Autofahren oder Hirnaktivitäten beim Betrachten von Verkehrssituationen erfährt.

Und ein besonders realitätsferner Test ist Bestandteil der *Medizinisch-Psychologischen Untersuchung* (= MPU, im Volksmund: „Idioten-Test“) zur Wiedererlangung der Fahrerlaubnis. Bei diesem sogenannten *Tachistoskopischen Verkehrsauffassungstest* (TAVT) werden für Sekundenbruchteile Fotos von Verkehrssituationen auf einem Monitor gezeigt. Nach jedem Bild wird per Multiple-Choice abgefragt, was auf dem Foto zu sehen war. Der Test soll den Umfang, die Schnelligkeit und die Genauigkeit der visuellen Wahrnehmung im Straßenverkehr erfassen. Doch er sagt praktisch nichts über die Leistungsfähigkeit eines Autofahrers im realen Straßenverkehr aus. Wenn er es täte, dann müssten Personen, die in diesem Test gute Ergebnisse erzielen, bei Fahrproben auch gut Auto fahren, und solche, die schlechte Testwerte erzielen, schlecht Auto fahren (= *Kriteriumsvalidität* oder *prognostische Validität*). Studien zeigen jedoch, dass der *Tachistoskopische Verkehrsauffassungstest* keine Vorhersagekraft für die Fahrleistung in einer Fahrprobe hat (Hannen et al., 1998; Held et al., 1993; Poschadel & Falkenstein, 2009). Das Beispiel zeigt: *Extern nicht valide* verkehrspsychologische Untersuchungsmethoden, die präzise Antworten auf Fragen liefern, die gar niemand gestellt hat und die am Kern der Sache vorbeigehen, sind nicht etwa nur eine Marotte von Forschern. Sie sind vielmehr schon seit Jahrzehnten bis in unseren Alltag vorgedrungen, haben sich dort festgesetzt und stellen sogar einen Teil der Datengrundlage für die Entscheidung dar, ob ein Mensch seinen Führerschein zurückbekommt oder nicht.

3. DER ZIELKONFLIKT ZWISCHEN REALITÄTSNÄHE UND METHODISCHER KONTROLLE

Die Gegenüberstellung der verschiedenen Methoden der Fahrerverhaltensbeobachtung zeigt, dass jede Methode ihre Stärken und Schwächen hat. Gleichzeitig gibt es einen Zielkonflikt: Je höher der Grad des realitätsnahen Fahrens ist, desto geringer ist der Grad an experimenteller Kontrolle – und umgekehrt (siehe auch Abbildung 1). Das eine Extrem bilden innerhalb der realitätsnahen Fahrsettings die *Naturalistic Driving Studies* (NDS), die dem normalen Fahren im Alltag sehr nahe kommen, aber den Nachteil besitzen, dass bei der Datenerhebung jede Art der Kontrolle fehlt und die Flut an aufgezeichneten Fahrdaten hinterher sehr schwer auszuwerten und zu interpretieren ist. Das andere Extrem bilden simulierte Fahrscenarien in einfachen Fahrsimulatoren oder Sitzkis-

ten, bei denen ein normales Autofahren mehr schlecht als recht imitiert wird. In den primitivsten dieser simulierten Szenarien (= rein visuelle Darbietung von Verkehrsszenarien) ist noch nicht einmal mehr eine Interaktion mit dem simulierten Fahrzeug oder Verkehr möglich. Tests mit solchen simulierten Szenarien erzielen zwar teilweise vorbildliche Werte bei wichtigen Testkriterien, d. h. sie sind *objektiv, reliabel, intern valide* und obendrein auch noch effizient und kostengünstig durchzuführen. Doch das ist nicht weiter erstaunlich, denn sie wurden auch genau auf diese Kriterien hin optimiert. Diese Vorteile sind jedoch mit einem gravierenden Nachteil erkauft: Diese Tests sind teilweise sehr realitätsfern und die Ergebnisse sagen nur wenig über das eigentliche Kriterium aus, nämlich wie ein Autofahrer sich im realen Straßenverkehr verhält.

Im Zielkonflikt zwischen Realitätsnähe und methodischer Kontrolle entscheidet sich die Mehrzahl der Forscher für die Kontrolle. Entsprechend sind auch Studien mit Fahrverhaltensbeobachtungen, die lediglich im Fahrsimulator durchgeführt wurden, in der Überzahl. Fahrsimulator-Studien sind Mainstream und längst zum Standard-Paradigma für Fahrverhaltensmessungen geworden. Leider führt dies auch dazu, dass das simulierte Fahren generell – einschließlich des erforderlichen technischen Mindestaufwands für eine realistische, immersive Simulation des Fahrerlebnisses – viel zu wenig kritisch hinterfragt wird. Die negative Konsequenz sind Simulator-Studien-Ergebnisse mit geringer Aussagekraft für das Fahren im echten Straßenverkehr.

Gegen die Kritik an Simulator-Studien mit mangelhafter *externer Validität* versucht sich die Szene der Fahrsimulator-Forscher mit einigen Standard-Argumenten zu verteidigen. Dies sind die fünf häufigsten Entschuldigungen:

Die Top 5 der Standard-Entschuldigungen für Fahrverhaltensstudien im Simulator mit geringer externer Validität:

1. „Es gibt ja auch viele (ältere) Studien, die mit noch viel einfacheren, unrealistischeren Fahrsimulationen gearbeitet haben. Verglichen damit ist unser Simulator wirklich sehr gut.“
2. „Es ist besser, die Untersuchung in diesem einfachen Standard-Simulator (bzw. mit dem *Lane Change Task*) zu machen, denn dadurch ist eine bessere Vergleichbarkeit mit anderen Untersuchungen gegeben, die das auch so gemacht haben.“
3. „Für diese konkrete, sehr spezielle, eingeschränkte Fragestellung reicht dieser einfache Simulator völlig aus.“
4. „Die zu messenden Effekte sind so klein, dass wir nur dann eine Chance haben, sie nachzuweisen, wenn wir unter hoch standardisierten, vollständig kontrollierten, experimentellen Bedingungen arbeiten und sämtliche Störfaktoren konstant halten oder eliminieren.“
5. „Wir wissen schon, dass unser Fahrsimulator nicht besonders realistisch ist. Aber wir können uns halt keinen High-End-Simulator für X Millionen leisten.“

Alle diese Argumente lassen sich leicht widerlegen:

Zu Argument 1, der *Vergleich nach unten*: Gewiss, es geht immer noch methodisch unsauberer. Aber methodisch fragwürdige Studien sind nicht der Maßstab, an dem man sich messen sollte. Und speziell zum simulierten Fahren ist das Maß aller Dinge die Wirklichkeit, also das Fahren in einem echten Auto im echten Verkehr, und nicht andere Fahrsimulationen, die noch weniger realistisch sind als die selbst verwendete.

Zu Argument 2, das *Vergleichbarkeits-Argument*: Hierbei handelt es sich um eine raffiniert verpackte Form des plumpen Arguments: „Das machen doch alle so“ bzw. „Das

haben wir schon immer so gemacht“. Dabei wird der Umstand ausgenutzt, dass Vergleichbarkeit und Replizierbarkeit von Untersuchungsergebnissen durch unabhängige Forschergruppen in der Wissenschaft tatsächlich ein wichtiger Aspekt ist (siehe [Teil 1](#)). Trotzdem darf das Ganze nicht dazu führen, dass man wider besseres Wissen an wenig geeigneten Forschungsmethoden festhält, wenn bessere zur Verfügung stehen.

Zu Argument 3, *Fokussierung auf einen ganz speziellen Untersuchungszweck*: Dieses Argument ist im Grunde gar keines. Es ist lediglich ein Beschwichtigen und Ablenken. „Die Kritik an unserer Fahrsimulation ist zwar grundsätzlich berechtigt, aber speziell in diesem Fall spielt das keine Rolle“. Verwiesen wird dann oft auf weitere Fehlerquellen in der Untersuchung (was das Ganze jedoch nur kritischer, aber nicht besser macht), wie z. B. dass der getestete System-Prototyp ohnehin völlig anders war als die endgültige Version im künftigen Serienfahrzeug. Zudem: Auch wenn man sehr spezielle Detailfragen untersucht, heißt das nicht, dass die Realitätsnähe eines Fahr Szenarios darauf keinen Einfluss hätte. Im richtigen Auto im realen Verkehr wären die Ergebnisse möglicherweise ganz anders.



Abbildung 8: Für Fahrstudien im Realfahrzeug gibt es mittlerweile smarte Komplettlösungen, so dass ein Kofferraum voller Messtechnik der Vergangenheit angehört. Auch die Aufzeichnung von Blickdaten oder physiologischen Daten ist damit kein Problem. Das Bild zeigt das Vehicle Testing Kit (VTK) der Firma Ergoneers. Mehr ist für die Durchführung eines Realfahrtstudie nicht nötig – abgesehen von einem Auto.

Zu Argument 4, *kleine Effektgrößen*: Dies ist ein Argument, das hauptsächlich von Wissenschaftlern vorgebracht wird, die sich gut mit Statistik auskennen. Es ist in der Sache auch richtig, aber es ignoriert die Frage der praktischen Relevanz eines nachgewiesenen kleinen Effekts. Ein Beispiel: Man kann unter hoch kontrollierten und standardisierten Bedingungen im Fahrsimulator unter Umständen durchaus nachweisen, dass mit einem bestimmten Assistenzsystem der Bremsweg um 10 Zentimeter kürzer ist als ohne dieses System – und zwar statistisch signifikant. Doch in der Praxis spielt ein 10 Zentimeter kürzerer Bremsweg keine Rolle. Statistisch signifikant ist eben nicht dasselbe wie bedeutsam, oder gar in der Praxis bedeutsam. Dieser Unterschied ist selbst vielen nicht bewusst, die eine statistische Ausbildung hinter sich haben. Wer Argument 4 benutzt, sollte sich daher selbstkritisch die Frage stellen, ob er nicht am eigentlichen Kern des Problems vorbei forscht.

Zu Argument 5, *hohe Kosten*: Sehr gute Fahrsimulatoren sind in der Tat extrem teuer. Und selbst sie sind immer noch weit vom Fahrerlebnis in einem richtigen Auto entfernt. Doch das Gegenargument lautet: Wieso überhaupt ein Fahrsimulator? Warum nicht gleich mit einem richtigen Auto in den realen Straßenverkehr gehen? Denn die Referenz ist die Realität selbst und nicht eine möglichst gute Simulation der Realität auf dem aktuellen Stand der Technik. Und anders als noch vor einigen Jahren ist eine solche Realfahrtstudie auch nicht mehr mit einem hohen technischen Aufwand und hohen Kosten verbunden.

4. FAZIT: EXPERIMENTELLE REALFAHRTSTUDIEN ALS VERNÜNFTIGER KOMPROMISS

Fahrstudien im Realfahrzeug erscheinen als vernünftiger Kompromiss, bei dem eine Balance zwischen realitätsnahem Autofahren einerseits und methodischer Kontrolle andererseits gewahrt bleibt. Die Vorteile überwiegen dabei deutlich:

Vorteile experimenteller Realfahrtstudien:

- Die Verhaltensbeobachtung findet bei einer normalen Fahrt in einem echten Auto und im echten Straßenverkehr statt. Alle Probleme mit unrealistisch simulierten Szenarien fallen weg.
- Die Datenerhebung ist im Gegensatz zu *Naturalistic Driving Studies* strukturiert, systematisch und hinsichtlich des Versuchsablaufs und Versuchsplans kontrolliert. Dadurch ist auch die Datenauswertung viel einfacher.
- Ein experimentelles Design ist ebenso möglich wie in einem Fahrsimulator, z. B. Versuchsbedingung mit und ohne Assistenzsystem.
- Über den CAN-Bus des Autos oder Zusatzgeräte wie *Mobileye* lassen sich die gleichen Fahrdaten aufzeichnen wie bei einer *Naturalistic Driving Study* oder in einem Fahrsimulator.
- Anders als bei einer *Naturalistic Driving Study* lässt sich das Blickverhalten mittels Eye-Tracking messen. Zwar sind hier Störeinflüsse (z. B. Sonnenlicht) größer als in einem Fahrsimulator unter Laborbedingung, doch dafür misst man das tatsächliche Blickverhalten im Straßenverkehr, das sich vom Blickverhalten in einem Simulator massiv unterscheidet.
- Der technische Aufwand ist viel geringer als bei einem Fahrsimulator.
- Die Kosten sind viel geringer als bei einer *Naturalistic Driving Study*, weil man nur ein einziges Fahrzeug mit Messtechnik ausstatten muss anstatt einer ganzen Fahrzeugflotte und weil der Erhebungszeitraum viel kürzer ist. Die Kosten sind auch geringer als bei einem (schlechten, billigen) Fahrsimulator, denn es ist nur die Messtechnik nötig – ein Auto hat jeder zur Verfügung.

Nachteile experimenteller Realfahrtstudien:

- Es gibt ein Unfallrisiko. Mit einem Fahrschulauto und einem mitfahrenden Fahrlehrer lässt jedoch auch dies minimieren.
- Es gibt durch die nicht vorhandene Kontrollierbarkeit des Fremdverkehrs – anders als im Fahrsimulator – mehr Freiheitsgrade beim Fahren und damit mehr Rauschen in den Daten. Es ist daher sinnvoll, eine Untersuchung mit einer größeren Versuchspersonenstichprobe als im Fahrsimulator zu machen, um die zu messenden Effekte auch statistisch absichern zu können.

Fahrverhaltensbeobachtungen in einem echten Auto im echten Verkehr sind also eine sinnvolle und praktikable Methode, die jedoch zu selten angewandt wird, obwohl – wie oben dargelegt – das simulierte Fahren unter Laborbedingungen große Defizite in punkto externer Validität hat. Die wichtigsten Argumente, mit denen Studien in realitätsfernen simulierten Fahrszenarien verteidigt werden, sind bei näherer Betrachtung nicht stichhaltig. Wer bereits einen Fahrsimulator besitzt und dafür viel Geld investiert hat, entscheidet sich eben gern für eine Fahrstudie im Simulator, selbst dann, wenn für die konkrete Fragestellung eine Realfahrtstudie besser geeignet wäre.

Doch anders als in vielen Bereichen der universitären Grundlagenforschung, bei denen eine praktische Anwendbarkeit ohnehin noch in weiter Ferne ist, lässt sich der Kritikpunkt der fehlenden Alltagsrelevanz bei Fahrverhaltensstudien nicht ignorieren. Wie sich ein Fahrer im Verkehr verhält, wie er beim Fahren Informationen verarbeitet und wie er Systeme bedient, sind durch und durch anwendungsbezogene und praxisrelevante Fragen.

Die Ergebnisse von Fahrstudien müssen deswegen in so realitätsnahen Settings gewonnen werden, dass sie sich auch auf das Fahren in einem realen Auto im realen Verkehr übertragen lassen. *Interne Validität* ist wichtiges, notwendiges Qualitätskriterium, reicht aber allein nicht aus. Nur wenn die Ergebnisse auf die Praxis übertragbar sind, sind die Schlussfolgerungen aus einer Fahrstudie korrekt und nützlich.

Ein Mehr an Untersuchungen im Realfahrzeug wäre daher wünschenswert. Die Technik dafür ist vorhanden und einfacher anzuwenden denn je. Denn es sollte nicht nur darum gehen, neuen Fahrerassistenz- und Fahrerinformationssystemen durch das Label „wissenschaftlich getestet“ einen Segen zu erteilen, ohne danach zu fragen, mit welchen Methoden eigentlich dieser Test durchgeführt wurde. Möchte man Fahrzeugtechnik entwickeln, die das Fahren tatsächlich komfortabler, stressfreier und sicherer macht, dann führt an methodisch sauberen Tests unter den realistischen, alltagsnahen Bedingungen einer Realfahrtstudie kein Weg vorbei.

5. LITERATUR

- Alliance of Automobile Manufacturers (2006). *Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems*, Washington, D.C.: Alliance of Automobile Manufacturers.
- Dagdelen, M., Reymond, G., Kemeny, A., Bordier, M. & Maizi, N. (2009). Model-based predictive motion cueing strategy for vehicle driving simulators. *Control Engineering Practice*, 17, 995-1003.
- Engström, J., Johansson, E. & Östlund, J. (2005). Effects of Visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8, 97-120.
- Guedry, F. E. (1970). Conflicting sensory orientation cues as a factor in motion sickness. *Fourth Symposium on the Role of the Vestibular Organs in Space Exploration* (S. 45-52). National Aeronautics and Space Administration, Washington DC 1970, NASA Report SP-187.
- Hannen, P., Hartje, W. & Skreczek, W. (1998). Beurteilung der Fahreignung nach Hirnschädigung: Neuropsychologische Diagnostik und Fahrprobe. *Nervenarzt*, 69, 864-872.
- Held, T., Lamberti, G. & Kubitzki, J. (1993). Psychose und Fahreignung – Entwicklung von Kriterien zur klinischen Beurteilung. *Die Rehabilitation*, 32, 155-161.
- ISO 26022 (2010). *Road vehicles. Ergonomic aspects of transport information and control systems. Simulated lane change test to assess in-vehicle secondary task demand*. British Standards Institution.
- Lietz, H., Petzoldt, T., Henning, M., Haupt, J., Wanielik, G., Krems, J., Mosebach, H., Schomerus, J., Baumann, M. & Noyer, U. (2011). *Methodische und technische Aspekte einer Naturalistic Driving Study*. FAT-Schriftenreihe 229, Forschungsvereinigung Automobiltechnik (FAT).
- Poschadel, S. & Falkenstein, M. (2009). *Testverfahren zur psychometrischen Leistungsprüfung der Fahreignung*. Berichte der Bundesanstalt für Straßenwesen: Mensch und Sicherheit, Heft M 203. Bremerhaven: Wirtschaftsverlag NW.
- Regan, E. C., Price, K. R. (1994). The frequency of occurrence and severity of side effects of immersion virtual reality. *Aviation Space and Environmental Medicine*, 65, 527-530.
- Straus, S. H. (2005). *New, improved, comprehensive, and automated driver's license test and vision screening system* (No. FHWA-AZ-04-559(1)). Phoenix: ESRA Consulting Corporation.
- Victor, T. W., Harbluk, J. L. & Engström, J. A. (2005). Sensitivity of eyemovement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8, 167-190.

Die Ergoneers GmbH wurde 2005 als Spin-off des Lehrstuhls für Ergonomie der Technischen Universität München gegründet. Heute ist das Unternehmen mit weltweiten Standorten in Deutschland (Hauptsitz bei München) und USA sowie zahlreichen Vertriebspartnern ein international wichtiger Partner für die Branchen Transport und Automotive, Marktforschung und Nutzerfreundlichkeit (Usability), Wissenschaft und Forschung sowie Sport und Biomechanik.

Neben der Entwicklung, Herstellung und dem Vertrieb von Mess- und Analysesystemen zur Erforschung von Verhalten und zur Optimierung der Interaktion von Mensch und Maschine bietet Ergoneers umfassende Kompetenz in allen Phasen des Studienablaufs. Zur Ergoneers-Produktpalette zählt vor allem die 360-Grad-Lösung D-LAB, eine umfassende Erfassungs- und Auswertungsplattform für Nutzer- und Verhaltensstudien, mit deren Software-Modulen sich Daten in den Bereichen Eye-Tracking, Data Stream, Video, Audio, Physiologie und CAN-Bus messen und analysieren lassen. Mit dem Ergoneers-Blickerfassungssystem Dikablis liefert Ergoneers zudem die passende Hardware, um professionelles Eye-Tracking im realen oder virtuellen Umfeld zu betreiben.

Ergoneers Group
Gewerbering 16
82544 Egling
Germany

T +49.8176.99894-0
F +49.8176.99894-15

Ergoneers of North America, Inc.
111 SW 5th Ave
Suite 3150
Portland, OR 97204
USA

T +1.503.444.3430

info@ergoneers.com
www.ergoneers.com

ERGONEERS
FROM SCIENCE TO INNOVATION