

Diagnostische Bewertung der Beherrschbarkeit von Fahrerassistenzsystemen

Patrick GALASKE¹, Ralf REISENAUER¹, Mehdi FARID¹, Klaus BENGLER²

¹ BMW Group

Knorrstraße 147, D-80788 München

² Lehrstuhl für Ergonomie, TU München

Boltzmannstraße 15, D-85747 München

Kurzfassung: Die fortschreitende Verfügbarkeit komplexer Fahrerassistenzsysteme in modernen Kraftfahrzeugen führt zu einer steigenden Bedeutung der Beherrschbarkeit dieser Systeme. Bestehende Ansätze zur Feststellung der Beherrschbarkeit besitzen jedoch keinen diagnostischen Charakter. Dieser Beitrag befasst sich mit der Entwicklung von Skalen zur Erfassung der Konstrukte Wahrnehmbarkeit der Kritikalität einer Fahrsituation und Durchführbarkeit der gewählten Gegenmaßnahme sowie deren relative Trennschärfe im Vergleich zur subjektiv wahrgenommenen Störung in einem Fahrsimulatorexperiment. Die Ergebnisse legen nahe, dass derartige Skalen durch eine Diagnostik der Nicht-Bherrschbarkeit einen Beitrag zur Entwicklung von Fahrerassistenzsystemen leisten können.

Schlüsselwörter: Fahrerassistenzsysteme, Beherrschbarkeit, Diagnose

1. Einleitung

Die Entwicklung von Fahrerassistenzsystemen (FAS) in modernen Kraftfahrzeugen hat zu einer Vielzahl von verfügbaren Assistenzsystemen geführt, deren Komplexität gleichzeitig wächst. Mit fortschreitender Entwicklung kommen diese Systeme immer häufiger zum Einsatz (Bosch-Gruppe, 2014). Dadurch gewinnt die Betrachtung der Beherrschbarkeit solcher Systeme in kritischen Fahrsituationen an Bedeutung.

Die Feststellung der Beherrschbarkeit von Fahrerassistenzsystemen (FAS) ist gemäß ISO 26262 ein notwendiger Bestandteil der Entwicklung solcher Systeme (ISO, 2011). Der Response 3 Code of Practice definiert die Beherrschbarkeit von FAS als ein Zusammenspiel der Fähigkeiten des Fahrers, die Kritikalität einer Fahrsituation wahrzunehmen, sich zu einer geeigneten Gegenmaßnahme zu entscheiden und diese Maßnahme erfolgreich durchzuführen (Response3, 2009).

Diese Definition des Konstrukts Beherrschbarkeit ist bisher für eine Bewertung von FAS nur begrenzt geeignet, da keine etablierte Operationalisierung der Unterkonstrukte existiert. Die empirische Bestimmung der Beherrschbarkeit von FAS fokussierte sich deswegen bislang auf objektiv prüfbare Kriterien der Nicht-Bherrschbarkeit (bspw. Simmermacher & Winner, 2011) sowie zusätzlich eindimensionale Bewertungen der subjektiven Störung (bspw. Neukum et al., 2008). Diese Ansätze sind jedoch ungeeignet, um die Ursache der Nichtbeherrschbarkeit zu ermitteln.

Dieser Beitrag befasst sich mit den Ergebnissen einer Simulatorstudie, die die Entwicklung und Erprobung eines Fragebogens zur Operationalisierung der Subkonstrukte der Beherrschbarkeit von FAS zum Ziel hat.

2. Methode

Die im Response 3 Code of Practice genannten Unterkonstrukte Wahrnehmbarkeit der Kritikalität der Situation, Entscheidung zu einer geeigneten Gegenmaßnahme und erfolgreiche Durchführung der Maßnahme besitzen, abhängig von der betrachteten Situation, zahlreiche Facetten, für die sich jeweils eine geeignete Skala entwickeln lässt. Eine Operationalisierung für eine spezielle Situation ist jedoch nicht das Ziel dieser Studie. Es soll vielmehr versucht werden, eine möglichst generische Skala zu entwickeln, die dazu geeignet ist, verschiedene Situationen bezüglich Ihrer Kritikalität entlang der verschiedenen Subdimensionen der Beherrschbarkeit von FAS zu bewerten.

Deswegen wurde mit einer Gruppe von Experten für jedes Teilkonstrukt ein Itemkatalog mit ca. je 30 Items erstellt, ohne einen Bezug zu einer bestimmten Situation zu erstellen. Diese wurden anschließend priorisiert und schließlich zu drei 6-Punkte-Likert-Skalen mit insgesamt 20 Items zusammengefasst. Die 6-Punkte-Skala wurde gewählt, um mittlere Urteile zu vermeiden. Die Skala wurde auf Basis der Empfehlungen in (Rohrmann, 1978) verankert. Dieses Entwicklungsverfahren wurde in Anlehnung an (Bortz & Döring, 2006) gewählt.

Diese situationsunabhängig entworfenen Skalen wurde in einem Simulatorversuch in einem der statischen Simulatoren des BMW Forschungs- und Innovationszentrums (FIZ) in einem gemischten between-within Design insgesamt 44 Probanden in 4 Situationen präsentiert. Dabei wurden die Wahrnehmbarkeit der Kritikalität sowie die Durchführbarkeit der Gegenmaßnahme manipuliert. Auf eine Manipulation der Schwierigkeit der Entscheidung zu einer Gegenmaßnahme wurde verzichtet, da dies selbst unter Simulatorbedingungen nur begrenzt möglich ist. Durch dieses Experiment sind 176 Bewertungen der beiden verbleibenden Konstrukte entstanden. Weiterhin wurden die subjektiv empfundene Störung der Fahrer mittels einer adaptierte Version der Störungsbewertungsskala (SBS) (Neukum et al. 2008) erhoben.

Zu Beginn aller 4 Situationen befand sich das Ego-Fahrzeug auf der rechten Spur einer Autobahn bei einer Geschwindigkeit von 80 km/h, die mittels Tempomat eingestellt wurde. Auf der linken Nebenspur befand sich dichter Verkehr, die rechte Spur war ein Standstreifen. Bei Beginn der kritischen Situation wechselte das Vorderfahrzeug auf die linke Spur und gab damit die Sicht auf ein stehendes Fahrzeug in der Ego-Spur frei. Gleichzeitig erschien auf dem Standstreifen eine Person, sodass ein Ausweichen des Ego-Fahrzeugs in beide Richtungen unmöglich war.

Die Wahrnehmbarkeit der Kritikalität der Situation wurde manipuliert, indem 50 Meter vor Beginn der kritischen Situation ein Warndreieck auf dem Standstreifen platziert wurde. Die Manipulation der Durchführbarkeit der Gegenmaßnahme bestand in einer Veränderung des Abstands zum Hindernis, bei dem das Vorderfahrzeug die Spur gewechselt hat. Bei einem späteren Spurwechsellvorgang des Vorderfahrzeugs verblieb dem Fahrer damit weniger Zeit, das Ego-Fahrzeug zum stehen zu bringen und damit eine Kollision zu vermeiden. Tabelle 1 stellt die Manipulation in den Fahrscenarien zusammenfassend dar.

Tabelle 1: *Tabelle der getesteten Szenarien.*

Situation	Warndreieck	Ausscherzeitpunkt Vorderfahrzeug
A	Vorhanden	Früh
B	Vorhanden	Spät
C	Fehlt	Früh
D	Fehlt	Spät

Für die Studie wurden 44 freiwillige Studienteilnehmer aus dem Umfeld des BMW FIZ rekrutiert. Das Durchschnittliche Alter der Teilnehmer lag bei 30 Jahren und die mittlere Fahrleistung lag bei ca. 10 000 km pro Jahr. Alle Teilnehmer wurden danach selektiert, einen gültigen Führerschein zu besitzen, jedoch keine erhöhte Fahrausbildung erfahren zu haben und nicht im Umfeld Fahrerassistenzsysteme zu arbeiten.

Die Teilnehmer wurden zufällig einer von vier Versuchsfahrten zugeordnet, die jeweils zwei kritische Situationen enthalten haben. Dabei wurde die Reihenfolge der Szenarien permutiert, um Reihenfolgeeffekte auszugleichen. Tabelle 2 stellt die Versuchsfahrten dar.

Tabelle 2: *Permutation der Szenarien in den Versuchsfahrten.*

Fahrt	Erste Situation	Zweite Situation
1	A	B
2	B	A
3	C	D
4	D	C

Die Teilnehmer haben nach einer Aufklärung über den Versuchszweck und einer kurzen Einführung in die genutzten Skalen eine Einführungsfahrt von 15 Minuten absolviert. Anschließend wurde die erste Prüfsituation gestartet. Nach dem ersten kritischen Szenario wurde zunächst der entworfene Fragebogen vom Probanden ausgefüllt und anschließend eine Bewertung auf der Störungsbewertungsskala abgefragt. Danach wurde die zweite Situation gestartet und dabei wie bei der ersten Situation vorgegangen.

3. Ergebnisse

Zunächst werden nun die Ergebnisse der Subskalen ausgewertet. Tabelle 3 stellt die Mittelwerte auf den Subskalen für die 4 Szenarien da. Hohe Skalenwerte bezeichnen schwierige Wahrnehmbarkeit und Durchführbarkeit bzw. große Störungen.

Tabelle 3: Mittelwerte der Subskalen in den kritischen Situationen.

Szenario	Skala Wahrnehmung [1-6]	Skala Durchführung [1-6]	SBS [0-10]
A	2,80	2,73	4,89
B	3,08	3,21	5,59
C	3,17	2,67	5,47
D	3,32	2,94	6,27

Der Mittelwert der Scores auf der Skala zur Wahrnehmbarkeit der Kritikalität liegen die Szenarien A und B erwartungsgemäß unter den Scores in den Szenarien C und D. Auch die Scores auf der Skala zur Durchführbarkeit der Gegenmaßnahme verhalten sich erwartungsgemäß ($B > A$ und $D > C$). Auf der Störungsbewertungsskala wurde die Verschärfung des Szenarios ebenfalls wiedergegeben, da auch hier der mittlere Score in A geringer liegt als in allen anderen Szenarien.

Anhand dieser Ergebnisse werden nun 4 Hypothesen überprüft. In den Szenarien mit Warndreieck (A & B) sei der Score auf der Skala zur Wahrnehmbarkeit der Kritikalität geringer als in den Szenarien ohne Warndreieck (C & D) (Hypothese 1). Dies gelte Ebenso für die Scores auf der Störungsbewertungsskala (Hypothese 2). In den Szenarien, in denen das Vorderfahrzeug spät ausschert (B & D) sei der Score auf der Skala zur Durchführbarkeit der Gegenmaßnahme höher als in den Szenarien, in denen das Vorderfahrzeug früh ausschert (A & C) (Hypothese 3). Dies gelte analog für die Scores auf der SBS (Hypothese 4).

Ein Lilliefors-Test ergab keine signifikante Abweichung der Scores von der Normalverteilung. Deswegen werden die Effekte der Manipulationen mittels einseitiger t-Tests überprüft. Für alle Signifikanztests gelte ein Signifikanzniveau von 5%. Tabelle 4 gibt einen Überblick über die Ergebnisse der Hypothesentests sowie die ermittelten Hedges-g Effektstärken.

Tabelle 4: Ergebnisse der Hypothesentests und Effektstärken.

Hypothese	p-Wert	Hedges-g Effektstärke
1	0,029*	0,42
2	0,059	(0,34)
3	0,021*	0,45
4	0,033*	0,41

Diese Ergebnisse legen nahe, dass die jeweiligen Manipulationen auf den Skalen der Wahrnehmbarkeit der Kritikalität sowie Durchführbarkeit der Gegenmaßnahme erfolgreich wiedergegeben wurden. Auf der Störungsbewertungsskala verursachte lediglich die Manipulation der Durchführbarkeit einen signifikanten Unterschied in den Scores. Die ermittelten Effektstärken liegen für alle signifikanten Hypothesen in der Nähe von 0,4. Sie können also als klein bis mittelgroß eingeordnet werden (Cohen, 1988).

4. Diskussion

Die in Kapitel 3 berichteten Ergebnisse zeigen exemplarisch, dass es möglich ist, situationsunabhängige Likert-Skalen zu entwickeln, mittels deren Unterschiede in der

Wahrnehmbarkeit der Kritikalität einer Fahrsituation sowie Durchführbarkeit der Gegenmaßnahme festgestellt werden können. Als wichtige Grundkonstrukte der Beherrschbarkeit von Fahrerassistenzsystemen können derartige Fragebögen einen wichtigen Beitrag zur Diagnose der Nicht-Bherrschbarkeit darstellen.

Der Vergleich der Effektstärken zwischen den erstellten Skalen sowie der SBS zeigt, dass die Werkzeuge die Manipulationen ähnlich scharf abbilden, wobei die Ursache des Unterschieds zwischen den Situationen anhand der SBS-Scores nicht ermittelt werden kann. Im Gegenzug ist der Einsatz der Störungsbewertungsskala mit deutlich weniger Aufwand verbunden.

Die hier dargestellten Ergebnisse wurden lediglich anhand eines Szenarios mit zwei Manipulationen erzeugt. Dadurch kann eine externe Validität nicht empirisch begründet werden. Da die beiden Skalen jedoch ohne Kenntnis der Prüfsituationen und unter dem Gesichtspunkt der Generalisierbarkeit erstellt worden sind, spricht aus theoretischer Sicht wenig dagegen. Gewissheit indes können nur zusätzliche Studien verschaffen.

Abschließend zeigen die Ergebnisse, dass ein standardisiertes Verfahren zur Diagnose der Nicht-Bherrschbarkeit von kritischen Fahrsituationen möglich ist und damit ein Beitrag zur Entwicklung von Fahrerassistenzsystemen geleistet werden kann. Ein Einsatz in einem breiteren Spektrum an Szenarien sowie eine Optimierung der Skalen auf Basis der erzielten Ergebnisse ist für die Zukunft wünschenswert. Ebenfalls könnte die Entwicklung eines Werkzeugs zur Messung der Schwierigkeit der Entscheidung zu einer Gegenmaßnahme einen weiteren Beitrag zur Diagnose der Nicht-Bherrschbarkeit von kritischen Fahrsituationen leisten.

5. Literatur

- Bortz, J., Döring, N. (2006). *Forschungsmethoden und Evaluation*. Berlin, Heidelberg: Springer.
- Bosch-Gruppe, (2014). *Bosch-Auswertung zu Fahrerassistenzsystemen*. Verfügbar unter: <http://www.bosch-presse.de/presseforum/details.htm?txtID=6962> (Abgerufen 5 Nov. 2014).
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York: Lawrence Erlbaum Assoc.
- ISO, International Organisation for Standardization (2011). *Road Vehicles - Functional Safety - Part 3: Concept Phase*. ISO 26262-3.
- Neukum, A., Lübbecke, T., Krüger, H.-P., Mayser, C. & Steinle, J. (2008) ACC-Stop&Go: Fahrerverhalten an funktionalen Systemgrenzen. In M. Maurer & C. Stiller (Hrsg.), 5. Workshop Fahrerassistenzsysteme - FAS 2008 141-150.
- Response3 (2009). *Code of Practice for the Design and Evaluation of ADAS*.
- Rohrmann, B. (1978). Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie*, 9(3) 222-245.
- Simmermacher, D., Winner, H. (2011). Beherrschbarkeit vom Gierstörungen durch ein Fahrerkollektiv. *Automobiltechnische Zeitschrift*, 113(9) 696-701.