



# Funktionale Anforderungen an Sensorik und Logik einer ATO-Einheit

Berichte  
des Deutschen Zentrums  
für Schienenverkehrsforschung

Bericht 66 (2025)





Berichte des Deutschen Zentrums  
für Schienenverkehrsforschung, Nr. 66 (2025)  
Projektnummer 2020-20-D-1202

# Funktionale Anforderungen an Sensorik und Logik einer ATO-Einheit

von

Birgit Milius, Baris Cogan  
Fachgebiet Bahnbetrieb und Infrastruktur, TU Berlin

Birte Thomas-Friedrich, Esther Bosch  
Institut für Verkehrssystemtechnik, Deutsches Zentrum für Luft- und Raumfahrt e.V.

Holger Last  
Siemens Mobility GmbH

Ulla Metzger, Dirk Leinhos  
Deutsche Bahn AG, DB Systemtechnik GmbH

im Auftrag des Deutschen Zentrums für Schienenverkehrsforschung beim Eisenbahn-Bundesamt

# Impressum

## HERAUSGEBER

Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt

August-Bebel-Straße 10  
01219 Dresden

[www.dzsf.bund.de](http://www.dzsf.bund.de)

## DURCHFÜHRUNG DER STUDIE

Technische Universität Berlin

Institut für Land- und Seeverkehr  
Fachgebiet Bahnbetrieb und Infrastruktur  
Salzufer 17-19, 10587 Berlin

## ABSCHLUSS DER STUDIE

05/2024

## REDAKTION

Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt  
Dr. Christian Klotz, Fachbereich Digitalisierung und Technik

## BILDNACHWEIS

Innentitel: Technische Universität Berlin

## PUBLIKATION ALS PDF

<http://www.dzsf.bund.de/veroeffentlichungen>

ISSN 2629-7973

doi: [10.48755/dzsf.250002.01](https://doi.org/10.48755/dzsf.250002.01)

Dresden, Mai 2025



This work is openly licensed via CC BY 4.0.

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.

# Inhaltsverzeichnis

<b>Kurzbeschreibung</b> .....	<b>8</b>
<b>Abstract</b> .....	<b>9</b>
<b>1 Einleitung</b> .....	<b>10</b>
<b>2 Projektziel und Vorgehensweise</b> .....	<b>11</b>
<b>3 Aufgaben von Triebfahrzeugführern</b> .....	<b>14</b>
3.1 Funktionale Beschreibung der Aufgaben von Triebfahrzeugführern .....	14
3.2 Definition von Standardaufgaben.....	14
<b>4 Einflussfaktoren auf die menschliche Leistungsfähigkeit</b> .....	<b>16</b>
<b>5 Sinne der Triebfahrzeugführer</b> .....	<b>18</b>
5.1 Sehen/Visuelle Wahrnehmung .....	18
5.1.1 Physiologische Einflussfaktoren auf die Leistung der visuellen Wahrnehmung.....	21
5.1.2 Individuelle Einflussfaktoren auf die Leistung der visuellen Wahrnehmung .....	24
5.1.3 Umweltfaktoren, die die Leistung der visuellen Wahrnehmung beeinflussen.....	26
5.2 Hören/ auditive Wahrnehmung.....	27
5.3 Gleichgewicht .....	28
5.4 Haptische und taktile Wahrnehmung/Tastsinn.....	29
5.5 Zusammenfassung.....	30
5.6 Anforderungen an die Wahrnehmungsleistung von Tf .....	31
<b>6 Ableitung von Szenarien</b> .....	<b>32</b>
6.1 Auswahl des zu untersuchenden Sinns .....	32
6.2 Auswahl einer Standardaufgabe .....	32
6.3 Unabhängige Variablen für die Simulatorstudie .....	34
6.3.1 Persönliche Einflussfaktoren .....	35
6.3.2 Betriebliche Parameter.....	36
6.3.3 Physikalische Eigenschaften des Reizes.....	36
6.4 Abhängige Variablen für die Simulatorstudie.....	40
<b>7 Experimentalstudien</b> .....	<b>41</b>
7.1 Versuchsumgebung.....	41
7.1.1 Fahrsimulator BBI.....	41
7.1.2 Fahrsimulator DLR .....	41
7.1.3 Onlinestudie.....	42
7.2 Szenarien.....	42

7.2.1	Szenarien BBI.....	43
7.2.2	Szenarien DLR.....	43
7.2.3	Szenarien der Online-Studie .....	43
7.3	Studiendesign .....	43
7.4	Forschungshypothesen .....	44
7.4.1	Simulatorstudien.....	45
7.4.2	Onlinestudie.....	45
7.5	Ablauf der Datenerhebung.....	46
7.6	Stichprobe .....	47
7.7	Datenvorverarbeitung.....	48
7.8	Statistische Auswertung.....	49
7.9	Definition des Regressionmodells .....	51
7.9.1	Simulatorstudien.....	51
7.9.2	Onlinestudie.....	52
<b>8</b>	<b>Ergebnisse Simulatorstudie BBI .....</b>	<b>54</b>
8.1	Deskriptive Analyse .....	54
8.2	Ergebnisse der Regressionsanalyse .....	55
8.2.1	Interpretation der Koeffizienten.....	58
8.3	Müdigkeit, Berufserfahrung, Alter, PVT .....	59
8.4	Zusammenerfassung der Ergebnisse .....	60
8.5	Semistrukturiertes Interview .....	61
<b>9</b>	<b>Ergebnisse Simulatorstudie DLR .....</b>	<b>63</b>
9.1	Deskriptive Analyse der Reaktionszeitdaten.....	63
9.2	Ergebnisse der Regressionsanalyse .....	65
9.2.1	Interpretation der Koeffizienten.....	66
9.3	Einfluss von Müdigkeit, Berufserfahrung, Alter .....	66
9.4	Zusammenfassung der Ergebnisse .....	67
<b>10</b>	<b>Ergebnisse Onlinestudie DLR.....</b>	<b>69</b>
10.1	Deskriptive Analyse der Reaktionszeitdaten.....	69
10.2	Ergebnisse der Regressionsanalyse .....	70
10.2.1	Interpretation der Koeffizienten.....	71
10.2.2	Einfluss von Müdigkeit, Geschlecht, Alter sowie Vorladen des Videos.....	72
10.3	Zusammenfassung der Ergebnisse .....	73
<b>11</b>	<b>Analyse der Gesamtstudie .....</b>	<b>74</b>
11.1	Vergleich der Simulatorstudien.....	74

---

11.2	Deskriptive Analyse .....	75
11.3	Mixed-Effects-Modell .....	79
11.3.1	Definition des Modells .....	79
11.3.2	Ergebnisse des Modells mit gemischten Effekten .....	80
11.3.3	Modellbewertung und -validierung .....	84
11.3.4	Vorhersage .....	86
11.3.5	Auswertungstool.....	88
11.4	Analyse der Reaktionszeit ohne Verwendung von Maßen der zentralen Tendenz .....	90
11.4.1	Vergleich der Verteilungen.....	91
11.4.2	Probabilistisches Modell.....	95
11.4.3	Ergebnisse der Überlebensanalyse.....	96
11.5	Zusammenfassung der Ergebnisse .....	98
11.6	Limitationen .....	99
<b>12</b>	<b>Diskussion .....</b>	<b>101</b>
<b>13</b>	<b>Abkürzungsverzeichnis .....</b>	<b>104</b>
<b>14</b>	<b>Abbildungsverzeichnis .....</b>	<b>105</b>
<b>15</b>	<b>Tabellenverzeichnis.....</b>	<b>108</b>
<b>16</b>	<b>Quellenverzeichnis .....</b>	<b>110</b>
<b>17</b>	<b>Anhänge.....</b>	<b>120</b>

# Kurzbeschreibung

Dieser Abschlussbericht des Projekts „Funktionale Anforderungen an Sensorik und Logik einer ATO-Einheit“ (ATO-SENSE) präsentiert die Ergebnisse der vier Arbeitspakete und diskutiert die Erkenntnisse ausführlich. Zunächst wurde eine funktionale Beschreibung der Aufgaben von Triebfahrzeugführern (Tf) durchgeführt. Basierend auf der detaillierten Aufgabenbeschreibung wurden Standardaufgaben der Tf definiert. Zudem wurden die Sinne des Menschen beschrieben, die für die Ausführung der Tf-Aufgaben notwendig sind und grundsätzlich auf ihren Einsatz im Rahmen der Aufgaben von Tf diskutiert. Im Rahmen einer Literaturrecherche wurde des Weiteren ein Überblick geschaffen, welche leistungsbeeinflussenden Faktoren für die Wahrnehmungsleistung dieser menschlichen Sinne in der Literatur bekannt sind und, wenn möglich, deren Bedeutung im Bahnbetrieb und beim Versagen von Funktionen analysiert. Diese leistungsbeeinflussenden Faktoren werden auch Performance Shaping Factors (PSF) genannt. Es wurde eine Expertenschätzung durchgeführt, um besonders relevante PSF für den Bahnbetrieb zu identifizieren. Aufbauend auf den erarbeiteten Ergebnissen wurden Experimente zur Bemessung der Wahrnehmungsleistung von Tf vorbereitet und durchgeführt. Die Szenarien für die Experimente wurden durch die Kombination einer Standardaufgabe mit ausgewählten Einflussfaktoren definiert. Die ausgewählte Aufgabe bestand aus einem Monitoring des Gleisbereichs und Betätigen des Signalhorns bei Auftauchen eines Objekts. Die ausgewählten Einflussfaktoren waren Kontrast und Größe des Objekts, Geschwindigkeit des Zuges und Zugsicherungssystem. Die Experimente wurden in einem Simulator am DLR und einem an der TU Berlin durchgeführt. Die Ergebnisse der Experimente wurden analysiert und kritisch diskutiert, wobei die Erkenntnisse aus der Literaturrecherche berücksichtigt wurden. Abschließend wurden die Ergebnisse der Experimente genutzt, um ein Modell zu entwickeln, das die Leistungsfähigkeit der visuellen Wahrnehmung von Tf in bestimmten Situationen repräsentiert. Die Analyse zeigte signifikante Effekte von Größe, Kontrast und Geschwindigkeitstyp auf die Reaktionszeiten. Die Koeffizienten von Größe, Kontrast und Geschwindigkeitstyp gingen in beiden Simulatoren in die gleiche Richtung und hatten einen ähnlichen relativen Einfluss. Dabei war der Einfluss der Größe mindestens doppelt so hoch wie der Einfluss von Kontrast. Die Ergebnisse dieser Studie tragen zur Erweiterung des Wissens über die Faktoren bei, die die RT beeinflussen, und bieten eine Grundlage für die Vorhersage von RT unter verschiedenen Bedingungen.

# Abstract

This final report of the project "Functional requirements for sensors and logic of an ATO-System" (ATO-SENSE) presents the outcomes of the work packages and provides a comprehensive discussion of the findings. Initially, a functional description of the tasks performed by train operators (Tf) was conducted. Based on the detailed task description, standard tasks of the drivers were identified. Additionally, the human senses necessary for executing Tf tasks were described and examined for their applicability within the context of train driver duties. A literature review was conducted to provide an overview of which performance-influencing factors for the perceptual performance of these human senses are known in the literature and, when possible, to analyze their importance in railway operations. An expert assessment was conducted to identify particularly relevant PSF for railway operations. Building upon the derived outcomes, experiments were prepared and conducted to assess the perceptual performance of train drivers. Scenarios for the experiments were defined by combining a standard task with selected influencing factors. The results of the experiments were analyzed and critically discussed, taking into account the insights from the literature review. Finally, the results of the experiments were employed to develop a model representing the visual perception performance of train drivers in specific situations. The analysis showed significant effects of size, contrast and speed type on reaction times. The coefficients of size, contrast and driving speed had an influence on the reaction times in the same direction in both simulator studies and had a similar relative influence. The effect of size on reaction times was at least twice as high as the effect of contrast. The results of this study contribute to the knowledge of the factors that influence RT and provide a basis for analysing reaction times under different conditions.

# 1 Einleitung

Digitalisierung und Automatisierung bringen grundlegende Veränderungen der Verkehrssysteme mit sich. Die Einführung des automatisierten Fahrens (ATO) eines Zuges wird den Bahnbetrieb in verschiedenen Ausbaustufen verändern. Diese Stufen wurden als Grades of Automation (GoA) im Nahverkehr eingestuft, sie haben sich aber auch für Hauptbahnen etabliert (DIN 62267:2010). So gibt es auch für Hauptbahnen bereits Pilotprojekte zum sogenannten Betrieb mit "ATO over ETCS" in GoA2, die Sicherheitsverantwortung liegt bei diesen Systemen jedoch weiterhin bei den Triebfahrzeugführern (Tf). Die GoA3/4-Spezifikation ist auf europäischer Ebene in der Entwicklung (X2Rail-1, 2019). Ab Automatisierungsstufe GoA3 wird die Sicherheitsverantwortung von Tf auf das System übertragen.

Unabhängig von dem gewählten Automatisierungsgrad muss für jedes neue oder signifikant geänderte System gezeigt werden, dass das damit verbundene Risiko vertretbar ist. Nach den gemeinsamen Sicherheitsmethoden (CSM) sollte eine Vertretbarkeit durch den Vergleich mit Referenzsystemen, durch Einhaltung eines anerkannten Regelwerks (code of practice) oder durch eine explizite Risikoanalyse nachgewiesen werden (Die Europäische Kommission, 2013). Neben der sicherheitlichen Betrachtung ist es aber auch wichtig, funktional zu verstehen, wie die Leistungsfähigkeit der heute eingesetzten menschlichen Tf beschrieben werden kann. Das Verständnis der Leistungsfähigkeit des Menschen und auch der Grenzen der menschlichen Leistungsfähigkeit ermöglichen eine Einschätzung der realistisch einzufordernden Leistung technischer Systeme, um sicherzustellen, dass neue Systeme mindestens so sicher, aber auch funktional vergleichbar sind, wie bisherige Systeme. Wesentlich für die erfolgreiche Aufgabenerfüllung ist, dass Tf ihre Sinne effektiv einsetzen. Im Rahmen der Tf-Aufgaben nutzen Menschen ihre Sinne beispielsweise zur Informationsbeschaffung, um zum einen die Außenwelt wahrzunehmen und zum anderen die korrekte Funktionsweise des Triebfahrzeugs zu überwachen. Die einzelnen Aufgaben der Tf erfordern somit sensorische und kognitive Anstrengungen in der Wahrnehmung von Informationen und deren Verarbeitung und Interpretation. Daher ist es wichtig, insbesondere die Leistungsfähigkeit der Tf im Hinblick auf die menschlichen Sinne zu verstehen, um Referenzanforderungen für automatisierte Systeme abzuleiten. Dieses Projekt soll zusammen mit dem Projekt zur Risikoakzeptanz für technische Systeme (ATO-Risk) einen Rahmen für das automatisierte Fahren definieren (Braband et al., 2023). In diesem Abschlussbericht werden die Methodik des ATO-Sense-Projekts sowie die Ergebnisse vorgestellt.

Im gesamten Bericht wird der Begriff "Triebfahrzeugführer (Tf)" gemäß den DB-Richtlinien verwendet. Der Begriff bezieht sich auf das Zugpersonal des EVU, für das die DB-Richtlinie 408.21-27 (Richtlinie 408.21-27:2014) gilt, und schließt explizit Tf jeden Geschlechts ein. Aus diesem Grund wird im Bericht auf eine weitere Verwendung geschlechtsspezifischer Begriffe verzichtet.

---

## 2 Projektziel und Vorgehensweise

Ziel des Projektes ist die Ermittlung eines Ansatzes zur Erfassung der Leistungsfähigkeit der menschlichen Sinne als Definitionsgrundlage der Anforderungen an ein technisches ATO-System. In diesem Projekt soll vorrangig analysiert werden,

- welche Sinne der Tf beim Steuern eines Zuges nutzt,
- welche Faktoren die Wahrnehmungsleistung dieser Sinne beeinflussen,
- wie schnell der Mensch auf einen sensorischen Input reagiert,
- und welche Bandbreite der Reaktionsgeschwindigkeit auf sensorischen Input zu erwarten ist.

Als Ergebnis des Projekts entstand unter Nutzung von im Projekt erhobenen Daten ein Modell, mit dem die Leistungsfähigkeit von Tf für spezifische Szenarien und Einflussfaktoren eingeschätzt werden kann. Inwieweit die Ergebnisse genutzt werden, um rechtlich bindende Vorgabewerte für zukünftige technische Systeme abzuleiten, ist nicht im Fokus des Projekts. Die notwendigen Vorarbeiten zur Erstellung des Modells erfolgten in insgesamt vier Arbeitspaketen. Im Folgenden wird das Vorgehen für die einzelnen Arbeitspakete detailliert beschrieben. Eine grafische Darstellung der Arbeitspakete ist in Abbildung 2 zu finden.

Das Ziel von Arbeitspaket (AP) 1 war die Erstellung von Vergleichsszenarien für die Analyse der Leistungsfähigkeit der menschlichen Sinne. Diese Vergleichsszenarien wurden im Rahmen von Simulatorexperimenten in AP 3 umgesetzt. Zur Definition der Vergleichsszenarien wurden zwei Arbeitsschritte durchgeführt sowie auf die Ergebnisse von AP 2 zurückgegriffen. Der erste Arbeitsschritt war die Identifikation der Aufgaben von Tf im konventionellen System. Für jede Aufgabe wurde definiert, welche menschlichen Sinne Tf bei der Ausführung der Aufgaben einsetzten. Es erfolgte eine Verdichtung der Aufgaben zu Aufgabenbereichen und Standardaufgaben unter Berücksichtigung vergleichbarer Sinnesnutzung und Zielorientierung. Zudem wurden als zweiter Arbeitsschritt unter Einbezug der Ergebnisse aus AP 2 interne und externe leistungsbeeinflussende Faktoren definiert, die die menschliche Leistungsfähigkeit beeinflussen. Die Bedeutung der identifizierten leistungsbeeinflussenden Faktoren für den Bahnbetrieb wurde analysiert und eine Expertenschätzung durchgeführt, um besonders relevante Faktoren zu identifizieren.

Ziel des AP 2 war die Erstellung einer Übersicht über den Stand der Forschung zur Leistungsfähigkeit der menschlichen Sinne, die Tf bei der Ausübung ihrer Aufgaben nutzen. Dazu wurde eine umfangreiche Literaturrecherche durchgeführt. Die Sinne wurden beschrieben, grundsätzlich auf ihren Einsatz im Rahmen der Aufgaben von Tf diskutiert und Einflussfaktoren auf die Leistungsfähigkeit der Sinne sowohl im Kontext des Zugsteuerns als auch in anderen relevanten Bereichen wie Straßenverkehr und Luftfahrt ermittelt. Für eine Beschreibung der Ergebnisse dieses Arbeitsschritts siehe Kapitel 4.

Im Mittelpunkt von AP 3 stand die Ableitung von Aussagen zur Leistungsfähigkeit des Menschen bei der Wahrnehmung. Es sollte im Rahmen von Experimenten festgestellt werden, wie schnell Tf unter Einfluss verschiedener leistungsbeeinflussender Faktoren über die Sinne Informationen wahrnehmen. Dabei wurde die Leistungsfähigkeit der menschlichen Wahrnehmung als die Geschwindigkeit einer Reaktion auf wahrzunehmende Informationen definiert (Reaktionszeit). Es zeigte sich, dass es zielführend war, sich für die Experimentalstudien auf den Sehsinn zu fokussieren. Zudem wurde aus den in AP 1 definierten Standardaufgaben nach mehreren Kriterien wie Automatisierbarkeit und Machbarkeit eine Standardaufgabe ausgewählt, die in der weiteren Untersuchung im Zentrum stehen sollte. Ausgehend von diesen Auswahlritten wurden durch eine Kombination der Standardaufgabe mit ausgewählten



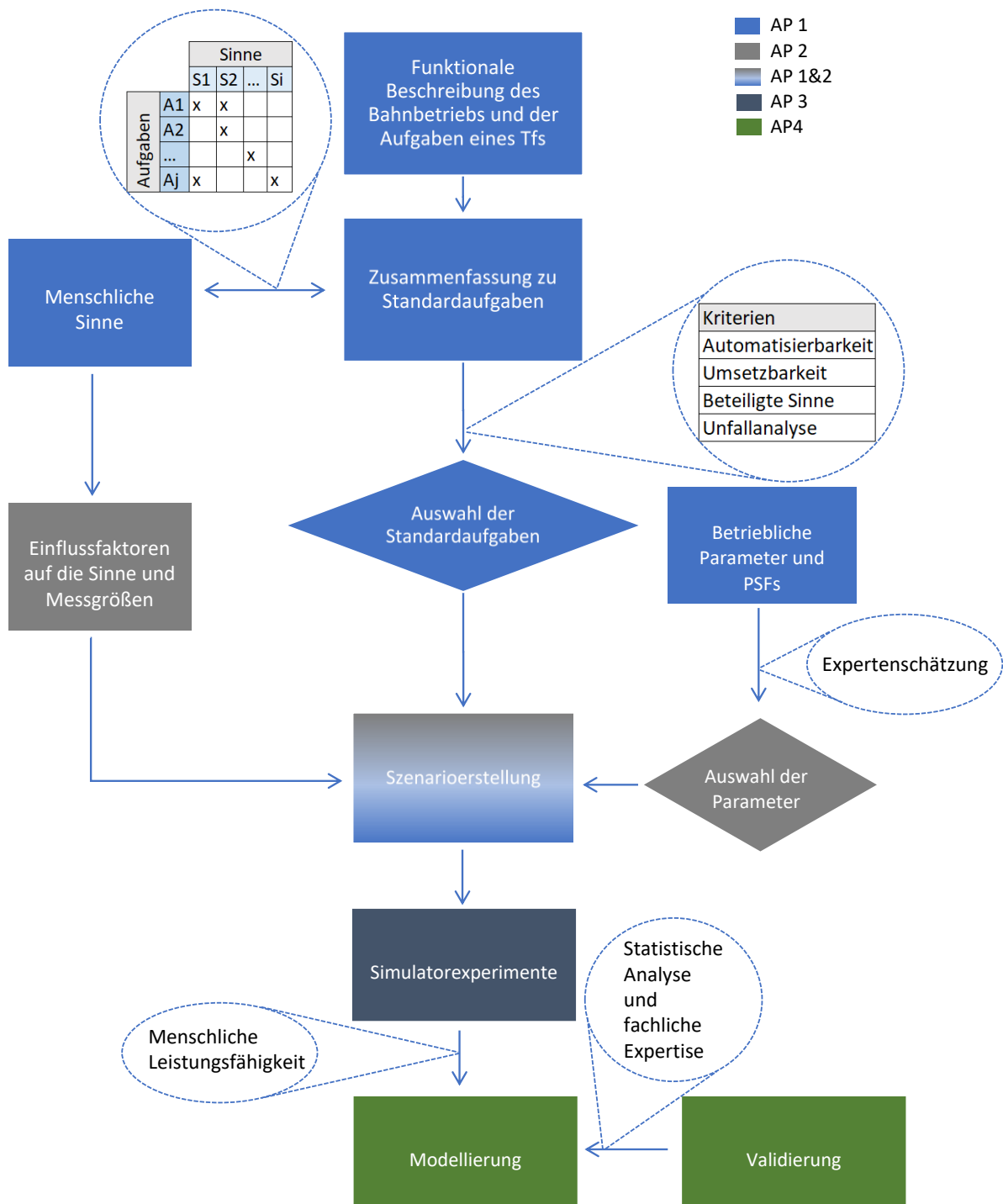


Abbildung 2: Vorgehensweise

## 3 Aufgaben von Triebfahrzeugführern

In diesem Kapitel wird die Identifikation und Kategorisierung der Aufgaben von Tf erläutert. Dazu wurde eine Analyse der Fahrdienstvorschrift vorgenommen und die extrahierten Aufgaben hinsichtlich von Aufgabenmerkmalen unterschieden. Der Prozess zur Definition von Standardaufgaben basierend auf Betriebsart, Ziel und Aufgabenbereich wird vorgestellt.

### 3.1 Funktionale Beschreibung der Aufgaben von Triebfahrzeugführern

Tf führen im Arbeitsalltag eine Vielzahl von Aufgaben aus. Das Projekt betrachtet die Tätigkeiten beim Steuern eines Zuges, so wie sie in der Modulgruppe 21-27 der Richtlinie 408 „Fahrdienstvorschrift“ der DB Netz AG (Richtlinie 408.21-27:2014) beschrieben werden. Die Richtlinie enthält Vorgaben und Schnittstellenregeln, die von allen Tf, die auf dem Netz der DB AG Züge fahren, beachtet werden müssen. Um die Aufgabenbeschreibung zu erstellen, wurden die Aufgaben und Verantwortlichkeiten der Tf in jedem Modul identifiziert und extrahiert. Nach der Extraktion der Einzelaufgaben wurden diese in einer Datentabelle zusammengestellt (Kapitel 17). Um die Schnittstelle zum Projekt ATO-Risk sicherzustellen, erfolgte zusätzlich eine Analyse der dort genutzten Richtlinie (DIN VDE V 0831-103:2020, Braband et al., 2023). Die dort genannten Aufgaben wurden denen aus der Fahrdienstvorschrift FV identifizierten gegenübergestellt. Aufgaben, die in beiden Richtlinien nicht enthalten sind, wurden nicht betrachtet.

Es wurde für jede Aufgabe festgehalten, welche Sinne die Tf bei der Ausführung dieser Aufgabe einsetzen. Berücksichtigt wurden dabei die Sinne Sehen, Hören, der Gleichgewichtssinn sowie der Tastsinn.

### 3.2 Definition von Standardaufgaben

Die identifizierten Aufgaben beschreiben den Arbeitsalltag eines Tf sehr kleinteilig. Es kann gezeigt werden, dass unterschiedliche Aufgaben vergleichbare Ziele verfolgen.

Aus diesem Grund wurden basierend auf der detaillierten Aufgabenbeschreibung Standardaufgaben definiert. Diese Standardaufgaben fassen die Einzelaufgaben der Tf so zusammen, dass die Tätigkeiten der Tf in ihrer Gesamtheit möglichst gut repräsentiert werden. Die Bestimmung der Standardaufgaben wurde in drei Schritten durchgeführt (Kapitel 17-Anhang 2).

- Im ersten Schritt wurden die Aufgaben der Tf in die Kategorien Regelbetrieb oder Störbetrieb eingeteilt, da davon auszugehen ist, dass hier unterschiedliche Anforderungen gestellt werden.
- Im zweiten Schritt wurden grundsätzliche Ziele (Aufgabenbereiche) wie z. B. „Auf Streckengeschwindigkeit beschleunigen“ oder „Am gewöhnlichen Halteplatz halten“ identifiziert.
- Vergleichbare Aufgabenbereiche wurden zu Standardaufgaben zusammengefasst. Die Standardaufgaben können Tabelle 1 entnommen werden.

TABELLE 1: STANDARDAUFGABEN

<b>Standardaufgabe</b>	<b>Aufgabenbereiche</b>
<b>Fahren/Fahrzeuggeschwindigkeit regeln</b>	Am gewöhnlichen Halteplatz halten
	Auf einem Bahnhof abfahren
	Beschleunigen
	Regelung der Geschwindigkeit
<b>Diagnose</b>	Erkennung und Feststellung von Unregelmäßigkeiten an Fahrzeugen oder Ladungen
	Ermittlung der Ursache eines unvorhergesehenen Anlasses
	Brand-, Rauch-, Entgleisungsdetektion
<b>Kommunikation Reisende</b>	Informieren der Reisenden über eine Notbremsüberbrückung
<b>Kommunikation betrieblich</b>	Meldung der Abreisebereitschaft an die Fahrdienstleiter (Fdl) in bestimmten Fällen
	Erteilung und Ausfertigung von Befehlen
<b>Kuppeln</b>	Beidrücken, Aufdrücken
<b>Instandsetzung</b>	Aussetzung der Fahrzeuge mit Mängeln aus dem Zug
	Reparatur/Instandsetzung von Lokomotiven und Wagen.
<b>Intervention im Störfall</b>	Reaktion bei Notruf
	Meldung der Störung an die Fdl/Betriebszentrale
	Evakuierung
<b>Überwachung Personenwechsel</b>	Türschluss
	Fahrgasttür-Überwachung
<b>Monitoring Fahrzeug</b>	Beachtung der Geschwindigkeiten nach Fahrplan und La
	Beobachtung die Signalinformation auf der Führer- raumanzeige
	Zugleitbetrieb
<b>Monitoring Infrastruktur</b>	Beobachtung ortsfesten Signals beim signalgeführten Zug
	Beobachtung der zu befahrende Strecke auf Unregelmäßigkeiten, die den Zug gefährden könnten

## 4 Einflussfaktoren auf die menschliche Leistungsfähigkeit

Die situativen Einflussgrößen, die Schwankungen in der menschlichen Leistung verursachen, werden als leistungsbeeinflussende Faktoren (Performance Shaping Factors, PSF) bezeichnet. Die EU-Richtlinie über Eisenbahnsicherheit unterstreicht die Bedeutung von menschlichen Einflüssen für die Aufrechterhaltung und Verbesserung der Eisenbahnsicherheit (EU directive 2016/798). Menschliche Einflüsse treten in vielen Phasen des Lebenszyklus von Eisenbahnsystemen auf, unter anderem in der Entwurfs-, Bau-, Betriebs- und Wartungsphase (Hammerl und Vanderhaegen, 2009).

PSF können verwendet werden, um eine aufgabenspezifische Zuverlässigkeit auf die positive oder negative Seite zu korrigieren, je nach der Situation, in der die Aufgabe ausgeführt wird (Lindner et al., 2014). Es existieren verschiedene PSF-Taxonomien, die allgemeinere oder detailliertere Faktoren enthalten.

Die Faktoren, die das menschliche Verhalten verbessern oder verschlechtern, werden häufig entsprechend ihrer Merkmale als intern und extern eingestuft. Die internen Faktoren beziehen sich auf individuelle Eigenschaften (z. B. Fähigkeiten, Motivationen, Erwartungen), während die externen Faktoren (z. B. Gerätekonstruktion, Verfahren und Anweisungen) mit den situativen oder umweltbedingten Umständen zusammenhängen (Swain und Guttman, 1983). Die THERP-Methode fügte eine dritte Klasse hinzu, die Stressfaktoren umfasst, die sich auf mentalen oder physischen Stress auswirken (z. B. Aufgabenbelastung, Ermüdung) (Boring, 2012). Nach dieser Klassifizierung resultieren psychologische Stressfaktoren (z. B. Ablenkungen) und physiologische Stressfaktoren (z. B. Müdigkeit und Unbehagen) aus einem Arbeitsumfeld, in dem die Anforderungen nicht den Fähigkeiten und Einschränkungen des Bedieners entsprechen.

Boring (2010) unterteilte die leistungsbeeinflussenden Faktoren anhand ihrer Messbarkeit in zwei Kategorien. Nach dieser Einteilung sind die direkten Faktoren messbar, während die indirekten Faktoren nur multivariat oder subjektiv bestimmt werden können.

Die R-PSF-Taxonomie basiert auf einer Literaturrecherche im Bereich menschlicher Faktoren und wurde anhand von Unfall- und Zwischenfallberichten validiert und von Fachexperten bewertet (Kyriakidis et al., 2012). Die Autoren befassten sich mit dreizehn PSF-Taxonomien in der Literatur aus den Bereichen Eisenbahn und anderer Verkehrsträger, Kernkraftwerk und Gesundheitswesen. Die PSF sind in sieben Kategorien unterteilt, nämlich statische persönliche, dynamische persönliche, Aufgaben-, Team-, Organisations-, Umwelt- und Systemfaktoren. Zeitabhängige PSFs, z. B. Müdigkeit oder Wetterbedingungen, werden als dynamische Faktoren bezeichnet,

Die deutsche Richtlinie zur menschlichen Zuverlässigkeit unterscheidet zwischen externen und internen leistungsbeeinflussenden Faktoren (VDI 4006). Externe Faktoren, die die Leistung beeinflussen, ergeben sich aus den technischen und organisatorischen Voraussetzungen. Die internen Faktoren ergeben sich aus der physiologischen und psychologischen Leistungsfähigkeit und -bereitschaft.

Rail-HRA, das in erster Linie für Tf entwickelt wurde, umfasst Methoden zur Identifizierung und Quantifizierung menschlicher Fehler (RSSB, 2005). Das Tool bietet eine Liste von Leistungsfaktoren, die die Zuverlässigkeit der menschlichen Leistung beeinflussen und verringern können. Das Railway Accident Reliability Assessment (RARA) ist der Nachfolger von Rail-HRA (RSSB, 2012). Eine Liste der fehlerverursa-

chenden Bedingungen, die die Aufgaben beeinflussen, ist enthalten. Die fehlerverursachenden Bedingungskategorien in der RARA sind wie folgt: Aufgabendesign, Schnittstellendesign, Kompetenzmanagement, Verfahren, Person, Umgebung.

Hammerl und Vanderhaegen (2009) haben die PSF aus den Taxonomien THERP und Rail-HRA auf Basis einer Ursache-Wirkungs-Beziehung für Tf und Fahrdienstleiter restrukturiert. Nach Ansicht der Autoren sind Begriffe wie Arbeitsbelastung, Stress und Vigilanz als passive abhängige Variablen bzw. als subjektive menschliche Reaktion auf die Einflussfaktoren zu verstehen (d. h. Vigilanz hängt von der Müdigkeit, Dienstplanung usw. ab).

Anhang 17.1 enthält einen Überblick über die identifizierten PSF. Die durchgeführte Literaturrecherche umfasst die Eisenbahn und andere Verkehrssektoren sowie die Bereiche Kernkraftwerke und Gesundheitswesen. Wie die Ausführungen zeigen, sind die PSFs umfassend und können ausgewählte menschliche Sinne oder die Leistung als Ganzes beeinflussen.

Basierend auf den grundlegenden Arbeiten zu PSF und den Erkenntnissen zu den Sinnen werden in Kapitel 5 die konkreten Faktoren identifiziert, die die sensorische Leistung von Tf beeinflussen.

## 5 Sinne der Triebfahrzeugführer

Die Leistungsfähigkeit der Tf wird unter anderem durch die Leistungsfähigkeit der menschlichen Sinne begrenzt. Als Grundlage für die folgenden Untersuchungen wurde eine Literaturrecherche zu den Sinnen der Tf durchgeführt. Für die Aufgaben der Tf wurden die folgenden Sinne als relevant eingestuft: Sehen (die visuelle Wahrnehmung); Hören (die auditive Wahrnehmung); der Tastsinn (die haptische und taktile Wahrnehmung) sowie der Gleichgewichtssinn. Die Sinne Schmecken und Riechen wurden in Abstimmung mit dem Auftraggeber nicht betrachtet. Im Folgenden sind die relevanten Ergebnisse für jeden dieser Sinne zusammenfassend dargestellt. Da sich bisherige Forschung im vergleichbaren Kontext hauptsächlich mit Straßenverkehr beschäftigt hat, bezieht sich ein großer Teil der im Folgenden genannten Quellen auf Kraftfahrzeugfahrer. Durch die Ähnlichkeit von Aufgaben, Anforderungen und Kontext argumentieren wir, dass diese Literatur eine geeignete Grundlage zur weiteren Forschung über Tf darstellt.

### 5.1 Sehen/Visuelle Wahrnehmung

Die visuelle Wahrnehmung bezeichnet die Aufnahme und Verarbeitung optischer Reize durch Licht. Dabei aktivieren die Lichtwellen beim Eintritt in das Auge die Rezeptorzellen auf der Netzhaut. Diese Rezeptorzellen lassen sich in Stäbchen und Zapfen unterteilen. Stäbchen ermöglichen ein Hell-Dunkel-Sehen, wodurch das Sehen auch bei Dunkelheit möglich ist (skotopisches Sehen). Zapfen sind verantwortlich für das Sehen bei ausreichender Helligkeit (photopisches Sehen) und ermöglichen die Wahrnehmung von verschiedenen Farben und Formen in scharfer Auflösung. Das für den Menschen sichtbare Farbspektrum liegt im Wellenlängenbereich zwischen ca. 380 nm und 770 nm (Höh et al., 2008). Bei Dämmerung sind sowohl die Stäbchen als auch die Zapfen an der visuellen Wahrnehmung beteiligt (mesopisches Sehen) (Becker-Carus und Wendt, 2016).

Stäbchen und Zapfen sind auf der Netzhaut unterschiedlich verteilt. Die Fovea ist ein Bereich im Zentrum der Netzhaut mit der höchsten Dichte an Zapfen. Aus diesem Grund erreicht der Mensch im Bereich der Fovea, also dem Zentrum der visuellen Wahrnehmung, das höchste Auflösungsvermögen und somit die größte Sehschärfe (Becker-Carus und Wendt, 2017, S. 87). Das Zentrum der visuellen Wahrnehmung entspricht einem Bereich von bis zu 3° Abweichung vom Fixationspunkt (Purves und Williams, 2004). In der Peripherie der Fovea, dem Bereich jenseits einer Abweichung von 3° vom Fixationspunkt, nimmt die Dichte der Zapfen ab und die Dichte der Stäbchen zu. Mit zunehmender Dichte der Stäbchen ist kein scharfes Sehen mehr möglich, aber die Sensitivität des Sehens wird größer („Sehen, dass etwas da ist“, Becker-Carus und Wendt, 2017, S. 87). Stäbchen sind äußerst sensibel für Bewegungen, Kontraste in Helligkeit und Farbe sowie plötzliche Veränderungen im peripheren Gesichtsfeld. Derartige Reize, die in der Peripherie auftauchen, haben somit ein erhöhtes Potential, die visuelle Aufmerksamkeit auf sich zu ziehen (Jonides, 1981, Lachenmayr, 2003, Lambert, Spencer und Mohindra, 1987, Yantis und Jonides, 1984, Wickens und Hollands, 1999, Remington, Johnston und Yantis, 1992). Am sogenannten „blinden Fleck“, wo der Sehnerv das Auge verlässt, ist keine Lichtverarbeitung möglich. Diese Lücke wird durch Verarbeitungsprozesse im Gehirn kompensiert (Goldstein, 2008, Becker-Carus und Wendt, 2017).

Es ist bedingt möglich, die Leistungsfähigkeit der menschlichen visuellen Wahrnehmung mit technischen Systemen zu vergleichen. Das DZSF-Projekt ATO-Sensorik erstellte eine Übersicht über verfügbare Sensorsysteme für ATO, die die Grundlage für spätere Datenerhebungsprojekte bilden kann (Leinhos, Flatt und Witte, 2022). Im Projekt wurden die horizontale und vertikale Winkelauflösung verschiedener derartiger Sensormodelle betrachtet. Die Winkelauflösungen der betrachteten LiDAR-Sensoren lagen innerhalb von 0,05°-0,94° und der RADAR-Sensoren innerhalb von 0,18°-3,30° (Leinhos, Flatt und

Witte, 2022). Während die Reichweite von LiDARs zwischen 50 m und 1000 m liegt, beträgt die Reichweite von RADARs 12,8 m bis 350 m. Im Vergleich dazu lässt sich die Leistungsfähigkeit der menschlichen visuellen Wahrnehmung wie folgt definieren: Nach der sogenannten Snellen-Definition ist das "normale Sehvermögen" die Fähigkeit des menschlichen Auges, einen einzelnen Strich mit der Größe von 1 Bogenminute (') zu erkennen (Walker, 2009). Dies entspricht dem Mindestauflösungswinkel des Auges, der es dem Menschen ermöglicht, kontrastreiche Striche, die 30 cm voneinander entfernt sind, in einer Entfernung von 1 km zu unterscheiden (Höh et al., 2008).

Einer der wichtigsten Aspekte von Fernerkundungsgeräten wie LiDAR und RADAR ist der große Öffnungswinkel, der bei einigen Sensoren bis zu 360° beträgt. Das zentrale Sichtfelds des menschlichen Auges wird im Gegensatz dazu als etwa 3° radial vom Fixationspunkt entfernt betrachtet, und die Größe des Nutzfeldes wird häufig mit 20° bezeichnet (Wolfe et al., 2017). Das menschliche Auge nutzt sakkadische Bewegungen, um einen größeren Bereich mit der hochauflösenden Fovea des Auges abzutasten. Die Auswahl des Sichtbereichs erfolgt auch durch die Bewegung des Kopfes und des Körpers. Nach DIN 16186-1 müssen Tf vom Führerstand aus mindestens ein Sichtfeld von 35° × 32,2° (H × V) haben (DIN EN 16186-1:2019).

In den nächsten Abschnitten werden Einflussfaktoren auf die Leistungsfähigkeit der visuellen Wahrnehmung (PSF) dargestellt. Diese Faktoren wurden in drei Kategorien betrachtet, nämlich physiologische Faktoren (Abschnitt 5.1.1), individuelle Faktoren (Abschnitt 5.1.2) und Umweltfaktoren (Abschnitt 5.1.3). Abbildung 3 fasst diese Einflussfaktoren und ihre Wechselwirkungen übersichtsartig zusammen.

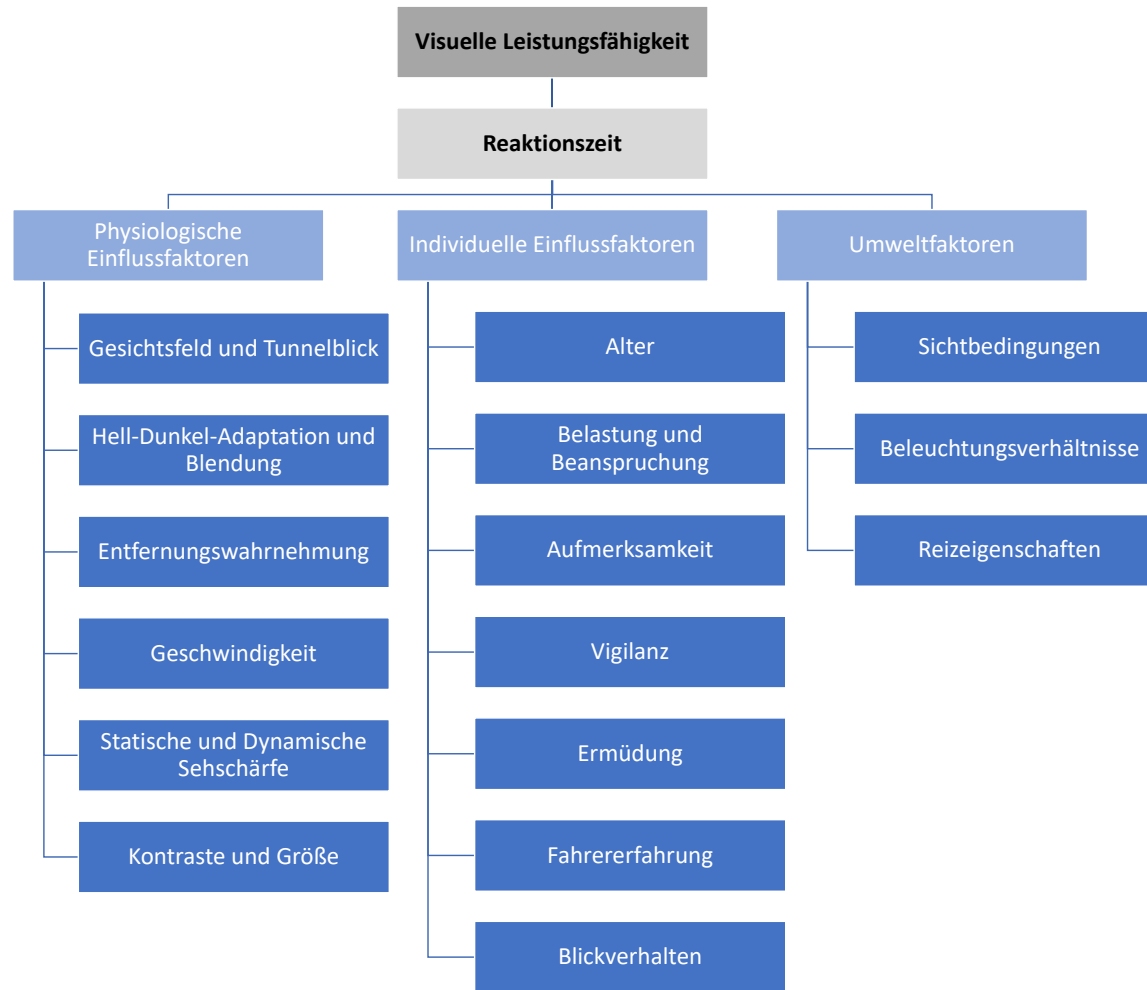


Abbildung 3: Einflussfaktoren auf die Leistungsfähigkeit der visuellen Wahrnehmung

## 5.1.1 Physiologische Einflussfaktoren auf die Leistung der visuellen Wahrnehmung

### 5.1.1.1 Gesichtsfeld und Tunnelblick

Mit dem Blickfeld wird der Bereich bezeichnet, in dem Objekte vom Menschen unter Zuhilfenahme von Augenbewegungen wahrgenommen werden können, ohne dass der Kopf dabei bewegt wird. Welche Position Objekte im Blickfeld einnehmen, kann über die Angabe von Winkeln kenntlich gemacht werden, wobei  $0^\circ$  dem Zentrum der Fovea entspricht. Das Gesichtsfeld, auch Useful Field of View (UFOV) genannt, stellt den Gesamtbereich dar, in dem ohne Zuhilfenahme von Augen- oder Kopfbewegungen visuelle Wahrnehmung möglich ist. Der genaue Bereich des UFOV ist individuell abhängig vom anatomischen Aufbau des Auges und seiner physiologischen Arbeitsweise. Zudem variiert er in Abhängigkeit verschiedener Faktoren, z. B. der jeweiligen Beanspruchung des Menschen und seines Alters (Schlag et al., 2009).

Erfolgt eine starke Einengung des UFOV, spricht man vom sogenannten Tunnelblick. Dabei gelingt die Objektwahrnehmung nur noch, wenn die Objekte nicht weit von der Blickrichtung abweichen. Verschiedene Faktoren können die Entstehung des Tunnelblicks fördern. Diese Faktoren sind hohe Geschwindigkeit, starke Ermüdung und Alkoholkonsum. Zudem wird der Tunnelblick gefördert, wenn Objekte oder Reize im peripheren Sichtfeld nur selten auftauchen, die wahrgenommen werden müssen (Land & Horwood, 1995; Schlag et al., 2009). Besonders Ermüdung scheint bei Tf von Bedeutung zu sein, da ihre Tätigkeit häufig im Schichtbetrieb und während der natürlichen Ruhephasen des menschlichen Körpers ausgeführt wird. Die Limitierung in der peripheren Wahrnehmung durch Müdigkeit könnte dann beispielsweise nachts oder bei Nebel dazu führen, dass Gefahrenquellen im peripheren Gesichtsfeld nicht registriert werden.

Für den Tf ist das periphere Gesichtsfeld aber für die Erkennung von Objekten am Rand der Strecke von Bedeutung, besonders im Bereich bis etwa  $30^\circ$ - $34^\circ$  Abstand vom Fixationspunkt (Lachenmayr, 2006). Objekte in diesem Bereich werden umso wahrscheinlicher fixiert, je näher sie sich am Zentrum der visuellen Wahrnehmung (Fovea) befinden, je auffälliger sie sind und je angemessener die Beanspruchung des Fahrers ist (Cohen, 1993). Helle, blinkende Lichter im peripheren Gesichtsfeld des Fahrers sind besonders gut geeignet, um die Aufmerksamkeit auf sich ziehen. Für Bewegungen im peripheren Gesichtsfeld gilt, dass Bewegungen umso besser detektiert werden, je schneller sie sind (Yeh und Multer, 2008).

### 5.1.1.2 Hell-Dunkel-Adaptation und Blendung

Hell-Dunkel-Adaptation bezeichnet die Anpassung des Auges an unterschiedliche Lichtverhältnisse. Für Tf ist dies zum Beispiel im Kontext von Tunnelfahrten oder grellem Licht relevant. Bei Dunkelheit schalten die Augen vom photopischen Sehen auf skotopisches Sehen um, d. h. es werden statt der Zapfen die Stäbchen für die visuelle Wahrnehmung genutzt. Bei Dämmerung findet das mesopische Sehen statt, d. h. es werden sowohl Zapfen als auch Stäbchen für die Wahrnehmung genutzt. Außerdem weiten sich die Pupillen, um den Lichteinfall zu maximieren. Es dauert etwa 10 Minuten für die Zapfen und etwa 30 Minuten für die Stäbchen, sich vollständig an Dunkelheit anzupassen (Goldstein, 1996). Da die Stäbchen eine höhere Lichtempfindlichkeit, aber eine geringere Sehschärfe erlauben, ist die Sehschärfe bei Dunkelheit auf etwa 5 % der zentralen Tagessehschärfe reduziert. Auch das Farbsehen ist stark eingeschränkt (Lachenmayr, 1995).

Die Umstellung von dunklen auf helle Lichtverhältnisse wird schon in wenigen Sekunden wirksam und ist bereits nach etwa 6 Minuten optimal. Falls sich die Leuchtdichte zu stark ändert, werden die Rezeptoren überbeansprucht und eine vorübergehende Blendung tritt ein, was auch „Adaptationsdefekt“ genannt wird. Dies tritt vorrangig bei Tunnelbeleuchtungen, bei Dämmerungsfahrten und bei Begegnungen von Fahrzeugen in der Dunkelheit auf (Hentschel, 2002). Eine physiologische Blendung hat eine Herabsetzung des Sehvermögens zur Folge (Gramberg-Danielsen, 1967).

Zusätzlich zur physiologischen Blendung gibt es das Phänomen der psychologischen Blendung. Dieses tritt ein, da Lichtquellen, welche eine Blendung hervorrufen können, die Aufmerksamkeit auf sich ziehen (Wördenweber, 2004). Die psychologische Blendung wird jedoch primär als „unangenehm“ bis „belästigend“ beschrieben und muss keine Herabsetzung des Sehvermögens zur Folge haben (Gramberg-Danielsen, 1967). Eine Einschränkung der Leistungsfähigkeit des Sehens bei der psychologischen Blendung besteht jedoch im Ermüden durch eine unbeabsichtigte sowie ständige Ablenkung der Blickrichtung, wobei diese zur Lichtquelle wandert. Diese Art der Blendung ist jedoch noch nicht ausgiebig erforscht (Schlag et al., 2009).

Die Empfindlichkeit für Blendung wird durch das Alter beeinflusst, wobei konkrete quantitative Angaben schwierig sind. Grundsätzlich können ältere Menschen deutlich blendungsempfindlicher sein als jüngere. Dies ist primär darin begründet, dass ältere Menschen Trübungen in den Augenteilen aufweisen, wodurch diese im Vergleich zu jüngeren Menschen empfindlicher für physiologische Blendungen sind (Ellinghaus, Schlag und Steinbrecher, 1990). Die gesteigerte Blendungsempfindlichkeit führt hauptsächlich zu einer Minderung der Sehschärfe, der Unterschiedsempfindlichkeit, des Form- bzw. Gestalterkennungsvermögens, der Tiefenwahrnehmung und der Wahrnehmungsgeschwindigkeit (Schlag et al., 2009).

### 5.1.1.3 Kontraste und Größe

Kontraste in der visuellen Wahrnehmung beschreiben den Unterschied zwischen Bereichen oder Objekten hinsichtlich ihrer Helligkeit oder Farbe. Die Größe des Kontrasts zwischen einem Objekt und dem Hintergrund spielt eine wesentliche Rolle bei der Frage, ob und wie gut das Objekt vom Menschen wahrgenommen werden kann. Die Grenze der Wahrnehmbarkeit zwischen Objekt und Hintergrund basierend auf dem Kontrast kann mithilfe des Weber'schen Gesetzes berechnet werden.

Das Weber'sche Gesetz  $k = \Delta R/R$  bezeichnet das Prinzip, dass sich zwei Reize um einen konstanten minimalen Prozentsatz – und nicht um einen konstanten Absolutbetrag – unterscheiden müssen, damit der Unterschied zwischen ihnen wahrgenommen wird (Stangl, 2021). Der gerade noch wahrnehmbare Unterschied  $\Delta R$  zum Vergleichsreiz  $R$  steht in einem bestimmten, gleichbleibenden Verhältnis  $k$ , der Weber-Konstante. Dabei ist das Weber'sche Gesetz nicht nur für die visuelle Wahrnehmung anwendbar, sondern für verschiedene Sinnesmodalitäten bei mittleren Reizintensitäten. In Bezug auf die visuelle Wahrnehmung beträgt die Weber-Konstante bei Helligkeit  $k = 0,017$  (Becker-Carus und Wendt, 2017). Bei geringer Helligkeitsintensität liegt die Wahrnehmungsschwelle für Kontraste höher, d. h. Reiz und Hintergrund müssen sich stärker unterscheiden, damit der Reiz wahrgenommen wird, als bei höherer Helligkeit.

Es ist allgemein anerkannt, dass die Stimulusintensität, wie z. B. Größe und Leuchtdichte, in einem umgekehrten Verhältnis zur Reaktionszeit steht. Es gibt jedoch verschiedene Ansichten über die genaue Natur dieser Beziehung. Das Piéronsche Prinzip besagt, dass die mittlere Reaktionszeit als Potenzgesetz mit zunehmender Leuchtdichte abnimmt (Piéron, 1914). Es wurde auch festgestellt, dass die mittlere

Reaktionszeit abnimmt, wenn Größe (oder die Fläche) des Reizes zunimmt (Bonnet, Gurlekian und Harris, 1992). Die Schätzung dieser Beziehung ist jedoch schwierig und hängt von der Bandbreite oder der Anzahl der Intensitäten ab (Pins und Bonnet, 1996). Dieser Effekt kann bei Vorhandensein anderer Faktoren, wie der Schwierigkeit der Aufgabe, unterschiedlich sein. So stellte (Bonnet, Gurlekian und Harris, 1992) beispielsweise fest, dass die Reaktionszeit als Funktion der Fläche eine gute Anpassung an das Pieronsche Gesetz darstellt, der Exponent der Funktion jedoch abnimmt, wenn die Leuchtdichte zunimmt. Es wurde festgestellt, dass die Auswirkungen nicht auf physikalische Reizeigenschaften beschränkt sind, sondern sogar durch Unterschiede in der wahrgenommenen Reizgröße moduliert werden können, die durch Entfernungswahrnehmung und visuelle Illusionen verursacht werden (Plewan und Rinkenauer, 2017).

#### 5.1.1.4 Entfernungswahrnehmung

Es besteht bei Menschen eine deutliche Divergenz zwischen objektiver und subjektiver wahrgenommener Distanz zu Objekten, wobei große Distanzen eher überschätzt werden. Auch bei der Geschwindigkeit gibt es eine Abweichung zwischen subjektiver und objektiver Geschwindigkeit, allerdings werden Geschwindigkeiten subjektiv eher unterschätzt, vor allem wenn längere Zeit mit einer hohen Geschwindigkeit gefahren wird (Hakkinen, 1963). Folglich wird subjektiv der Raum zwischen der eigenen Person und einem sich nähernden Objekt überschätzt, während die Annäherungsgeschwindigkeiten unterschätzt werden.

Wenn die Sicht eingeschränkt ist, wie z. B. bei Nebel, ist aufgrund der geringen Kontraste die Erkennung von Texturgradienten verschlechtert (Richter und Schlag, 2000). Die Fixation wird in den Nahbereich verlagert, wodurch eine andere Sinneswahrnehmung und ein anderes Erleben der Verkehrssituation entstehen. Dabei wird die Entfernung zu einem Objekt als zu 60 % größer eingeschätzt als sie tatsächlich ist (Cavallo und Cohen, 2001).

#### 5.1.1.5 Geschwindigkeit

Die Blicktiefe und Blickfeldweite korrelieren mit der eigenen Geschwindigkeit (OECD / ECMT, 2006, Weller et al., 2006). Je höher die eigene Geschwindigkeit ist, desto ferner ist die visuelle Orientierung und desto enger ist das Gesichtsfeld. Menschen in Bewegung blicken ca. drei Sekunden voraus, weswegen der Fixationspunkt vorverlagert wird, was in einer Verschlechterung der peripheren Wahrnehmung im Nahbereich resultiert (Land und Horwood, 1995). Mit zunehmender Geschwindigkeit verlagern Fahrer verstärkt ihren Fixationspunkt zu voraus liegenden Informationen, was bei ungünstigen Sichtbedingungen zu Problemen führt (Land und Horwood, 1995). Die Verengung der peripheren Sehleistung hängt jedoch nicht nur von der Geschwindigkeit, sondern auch von der Komplexität der Umgebung ab (Miura 1990, Recarte und Nunes, 2000, Weller, 2006).

Das Erkennen von bewegten Objekten wird durch die dynamische Sehschärfe ermöglicht. Diese nimmt bei zunehmender Geschwindigkeit zu, bevor sie ein Maximum erreicht und danach kontinuierlich abfällt. Diese Sehfunktion kann durch hohe Beleuchtungsstärken stark verbessert werden (Eckert, 1993).

Zudem beeinflusst die Fahrgeschwindigkeit die visuelle Wahrnehmung, da sie das Blickverhalten verändert. Bei einer simulierten Autofahraufgabe führt eine höhere Fahrgeschwindigkeit zu schnelleren Reaktionen auf Fahrbahnmarkierungen (Cao und Wang, 2004), aber zu einer Verengung des nutzbaren Sichtfeldes (Rogé et al., 2004) (Rogé et al., 2004). Ein Einfluss der Geschwindigkeit auf die visuelle Leistungs-

fähigkeit konnte auch in Bezug auf das Blickverhalten bei Tf nachgewiesen werden. In einer Studie zeigten sich bei höheren Geschwindigkeiten mehr vertikale und weniger horizontale Blickfixationen, bei niedriger Geschwindigkeit hingegen mehr horizontale Blickbewegungen und der Blick wurde von Seite zu Seite schweifen gelassen (Suzuki, Yamauchi und Matsuura, 2019).

## 5.1.2 Individuelle Einflussfaktoren auf die Leistung der visuellen Wahrnehmung

### 5.1.2.1 Alter

Auch wenn keine pathologischen Veränderungen im Auge vorliegen, nehmen die Fähigkeiten der visuellen Wahrnehmung durch physiologische und optische Veränderungen des Auges mit dem Alter ab. Besonders ab dem Alter von etwa 70 Jahren ist die Sehfähigkeit stark eingeschränkt (HaegerstromPortnoy, Schneck und Brabyn, 1999). Das Ausmaß der Einschränkung der Sehfähigkeiten ist jedoch interindividuell unterschiedlich. Daher ist das Alter kein ausreichender Indikator, um die Sehfähigkeit einer Person allein damit zu bewerten.

Eine der Sehfähigkeiten, die mit dem Alter abnimmt, ist die Fähigkeit zur Akkommodation. Bei der Akkommodation wird die Form der Linse und damit die Brechkraft des Auges verändert, wodurch eine dynamische Anpassung an die Entfernung von Objekten ermöglicht wird, sodass diese scharf gesehen werden können (Goldstein, 2008, Becker-Carus und Wendt, 2017). Die Abnahme der Fähigkeit zur Akkommodation resultiert in einer fortschreitenden Weitsichtigkeit, die bereits mit etwa 20 Jahren beginnt (Becker-Carus und Wendt, 2017). Auch die Lichtsensitivität nimmt etwa ab dem 20. Lebensjahr relativ konstant ab, wodurch die Sehschärfe in lichtarmer Umgebung verschlechtert wird.

Auch das räumliche Sehen bei Blendung, reduziertem Licht oder Kontrast ist im Alter deutlich eingeschränkt. Zudem sind mit zunehmendem Alter während des Zustands geteilter Aufmerksamkeit die Tiefenwahrnehmung, die Farbunterscheidung und die periphere Sicht ebenfalls eingeschränkt (Haegerstrom-Portnoy, Schneck und Brabyn, 1999). Die Tagessehschärfe bei 100 % Kontrast, das periphere Sehen und die selektive Aufmerksamkeit nehmen ebenfalls im Alter ab. Bei der visuellen Suche verlängern sich die Reaktionszeiten und es kommt häufiger zu Auslassern, also Objekten, die nicht entdeckt werden (Schlag und Beckmann, 2013).

### 5.1.2.2 Belastung und Beanspruchung

Fahrzeugführer können nur selektiv Informationen aufnehmen und müssen dabei Informationen mit Verkehrsrelevanz herausfiltern. Miura (1992) zeigt, dass sich in gefährlichen Situationen mit hoher Informationsdichte die Erkennungszeit wichtiger Objekte verlängert, da die Verarbeitungskapazität begrenzt ist. Eine weitere Sinnesgrenze besteht darin, dass der Fahrzeugführer im Durchschnitt nur 3 Objekte pro Sekunde fixieren kann (Schlag und Heger, 2002); (Cohen, 1987) (Cohen, 1993).

Auch die Größe des Gesichtsfelds (UFOV), also der Bereich, in dem visuelle Informationen ohne Augenbewegungen wahrgenommen werden können, ist abhängig von der Beanspruchung. Ermüdung sowie Unter- und Überbeanspruchung des Fahrzeugführers beeinflussen das UFOV negativ, indem sich der Bereich des UFOV verengt. Bei hoher Beanspruchung sind besonders die begrenzten mentalen und visuellen Verarbeitungskapazitäten für die Verringerung des UFOV entscheidend, bei niedriger Beanspruchung verringert das geringe Aufmerksamkeitsniveau das UFOV (Cohen, 1987, Schlag& Heger, 2000).

Insbesondere die Beeinflussung des UFOV durch eine Unterbeanspruchung ist im Kontext der Tf-Aufgaben relevant. Studien haben gezeigt, dass das Fahren der Tf unter höheren Automatisierungsstufen zu einer starken Unterbeanspruchung führt (Brandenburger et al., 2018). Allerdings muss berücksichtigt werden, dass die aus Aufgaben resultierende Beanspruchung sich interindividuell unterscheidet. Zur Erfassung der Beanspruchung eignen sich Leistungsmaße, Selbstauskünfte (Richter und Hacker, 1998) oder physiologische Parameter (Kramer, 1990).

Auch die aus einer Nebenaufgabe resultierende Beanspruchung kann sich negativ auf die Gefahrenerkennung auswirken. So zeigte sich bei Kraftfahrern in einer Untersuchung von Borowsky et al. (2016) nach einer zwei Sekunden langen Unterbrechung der Fahraufgabe eine schlechtere Erkennungsleistung einer darauffolgenden Gefahrenstelle, wenn diese Unterbrechung eine komplexe Nebenaufgabe beinhaltete.

### 5.1.2.3 Ermüdung

Aus erhöhter Müdigkeit resultieren Beeinträchtigungen der visuellen Wahrnehmung. So gehen mit zunehmender Müdigkeit beispielsweise eine Einengung des Gesichtsfeldes und die Abnahme der Akkomodationsfähigkeit einher (Schlag et al., 2009). Weiterhin mehren sich auch Wahrnehmungsfehler, Schwächen in der Hand-Auge-Koordination und eine verringerte Aufmerksamkeitsleistung (Schlag et al., 2009). Insbesondere eine verringerte Aufmerksamkeitsleistung kann sich dabei auf die Verkehrssicherheit auswirken. So zeigte sich in mehreren retrospektiven Untersuchungen zu sicherheitskritischen Vorfällen im Bahnbereich, dass diesen sicherheitskritischen Vorfällen häufig Aufmerksamkeitsfehler vorausgingen (Baysari et al., 2008; Edkins & Pollock, 1997; Kyriakidis et al., 2012; van der Flier & Schoonman, 1988).

Eine erhöhte Müdigkeit geht weiterhin einher mit einer reduzierten Vigilanz. Vigilanz ist das Aufrechterhalten der Aufmerksamkeit über einen längeren Zeitraum in einer monotonen Umgebung mit einer geringen Anzahl an auftretenden Reizen (Schmidt et al., 2009). Eine reduzierte Vigilanz wirkt sich auf die Sinnesleistung, die Aufmerksamkeit, die Reaktionszeit und die Verarbeitung sowie den Zusammenschluss verkehrsrelevanter Informationen aus (Sprenger, 2001)(Sprenger, 2001). Die Vigilanzleistung weist somit Relevanz für die Verkehrssicherheit auf. In einer Untersuchung konnte gezeigt werden, dass Tf, die in der Vergangenheit an Unfällen ursächlich beteiligt waren, schlechtere Vigilanzleistungen zeigten, als unfallfreie Tf (Hani Tabai et al., 2018). Weiterhin zeigte sich bei Schmidt et al. (2009), dass Kraftfahrer ihre Vigilanzleistung trotz messbarer Verschlechterungen, beispielsweise in der Reaktionszeit, subjektiv überschätzen.

Eine Studie mit 26 Tf zeigte, dass sowohl die subjektive Bewertung als auch physiologische Maße (gemessen mit Elektroenzephalographie und Elektrokardiogramm) für Müdigkeit während einer zweistündigen Simulatorfahrt signifikant zunahm (Brandenburger, Wittkowski und Naumann, 2017). In einer Untersuchung mit Kraftfahrzeugführern wurde festgestellt, dass in der Zeit zwischen 12 Uhr und 15 Uhr sowie zwischen 23 Uhr und 5 Uhr am häufigsten ein Einschlafen am Steuer auftritt. Dies ist bedingt durch die circadiane Rhythmik (Kuratorium für Verkehrssicherheit, 2006). Bei den Unfallzahlen ist dabei erkennbar, dass sich die besonders hohen Raten während der Nacht mit den Erholungs- bzw. Ermüdungsphasen decken (Heinz, 1982). Symptome von Ermüdung werden durch Eintönigkeit der Fahrbedingungen sowie durch wenige Reize in der Umgebung verstärkt (Lachenmayr, 1995). Zudem ist es problematisch, dass Fahrer häufig ihre eigene Wachheit und Fahrtüchtigkeit nicht korrekt einschätzen können (Gründl, 2005).

#### 5.1.2.4 Fahrerfahrung

Die Fahrerfahrung bestimmt maßgeblich, welche Objekte für einen Fahrzeugführer in verschiedenen Verkehrsumwelten und -situationen handlungsleitend sind (Schlag, 2008). Dies geschieht unter anderem mithilfe subjektiver, erfahrungsbasierter Kategorien bezüglich der Straßenumwelt, die in bestimmten Situationen zur Aktivierung angemessen wirkender Verhaltensmuster führen. Die daraus resultierenden Erwartungen des Fahrers beeinflussen die Wahrnehmung, auch im Hinblick auf deren Selektivität und Akzentuierung (Schlag und Heger, 2002). Auch die Erkennungs- und Reaktionszeit werden durch gebildete Erwartungen beeinflusst. Beispielsweise zeigte sich, dass der Abstand, mit dem ein Objekt erkannt wird, bei unerwarteten Objekten im Vergleich zu erwarteten Objekten nur die Hälfte beträgt, unerwartete Objekte also erst später erkannt werden (Shinar, 1985). Der Einfluss der Erwartungen geht dabei so weit, dass Bartmann et al. (1993) feststellten, dass deren positive Effekte, insbesondere die Antizipation von verkehrsrelevanten Informationen, teilweise Sehdefizite kompensieren kann.

Weiterhin wirkt sich die Fahrerfahrung auf die Blicksteuerung aus (Cohen, 1984, zitiert nach Cohen, 1993, Cohen, 2008). Dies äußert sich beispielsweise darin, dass Fahranfänger im Automobilkontext im Vergleich zu erfahrenen Fahrern seltener Fixierungen in der Ferne zeigen, größere Blickbewegungsamplituden aufweisen und ihre Orientierung an weniger verkehrsrelevanten Umgebungsinformationen ausrichten (Mourant und Rockwell, 1972). Laut Crundall und Underwood (1998) werden diese Unterschiede und die daraus resultierenden Filtereffekte allerdings hauptsächlich in komplexen Situationen bzw. bei hoher Beanspruchung relevant für die Verkehrssicherheit. Denn insbesondere in diesen Situationen erhöht sich die Unfallgefahr, wenn die Situationserfassung und Gefahrenantizipation nur unzureichend funktionieren. Darüber hinaus erschwert das ungünstigere Blickverhalten bzw. die geringere Erfahrung der Fahranfänger teilweise die Aufnahme verkehrsrelevanter Informationen und verringert so die Entscheidungsgeschwindigkeit (Underwood et al., 2002). Erfahrene Fahrer können daher durch ein konzeptgesteuertes visuelles Suchverhalten sowohl kognitive Ressourcen schonen als auch schnellere Entscheidungen treffen.

Ein positiver Einfluss der Fahrerfahrung bzw. in diesem Fall Berufserfahrung konnte auch auf das Blickverhalten von Tf nachgewiesen werden. In einer Studie von Horiguchi et al. (2015) zeigte eine Gruppe von erfahrenen Tf im Vergleich zur Gruppe unerfahrener Tf kürzere und mehr Blickfixationen, was auf effizientere Blickverteilungsmuster hindeutete. Die unerfahrenen Tf fixierten dabei häufiger und länger auch kabineninterne Bereiche wie beispielsweise das Tachometer und waren so schlechter in der Lage Informationen aus der Umgebung zu extrahieren (Horiguchi et al., 2015).

#### 5.1.3 Umweltfaktoren, die die Leistung der visuellen Wahrnehmung beeinflussen

Dämmerung, Dunkelheit und schlechte Sichtbedingungen schränken die Leistung der visuellen Wahrnehmung ein. Die Dämmerung erfordert eine hohe Anpassungsleistung des visuellen Sinns. Einerseits ändern sich die Beleuchtungsverhältnisse beim Übergang vom natürlichen zum künstlichen Licht durch Straßenlaternen bzw. Scheinwerfer der Fahrzeuge, wodurch sich in dieser Phase sowohl die Kontraste als auch das Leuchtdichteniveau relativ schnell ändern. Andererseits entsteht nach Sonnenuntergang durch das Streulicht des Himmels eine gleichmäßig diffuse Illuminierung (Eckert, 1993). Bei Dunkelheit verringern sich sowohl die Kontraste als auch die Wahrnehmbarkeit vieler Objekte. Die Kontrastminderung resultiert dabei aus der räumlich begrenzten Belichtung der nächtlichen Lichtquellen. Die verringerte Wahrnehmbarkeit entsteht, da häufig die absolute Wahrnehmungsschwelle erreicht wird (Schmidt-Clausen und Freiding, 2004).

Wichtig ist zu beachten, dass trotz der Informationsdefizite bei Nacht viele Fahrzeugführer ihre Sehfähigkeiten überschätzen und dabei das entfernteste beleuchtete oder selbstleuchtende Objekt, welches sie wahrnehmen können, mit der jeweiligen Sichtdistanz gleichsetzen (Cohen, 1993). Die Sichtdistanz für beleuchtete oder selbstleuchtende Objekte entspricht aber nicht der Sichtdistanz, mit der unbeleuchtete Objekte wahrgenommen werden können, da diese Sichtdistanz deutlich kürzer ist.

Weiterhin wirken sich die Sichtbedingungen auch auf das Blickverhalten aus. Beispielsweise zeigten sich in einer Untersuchung mit Kraftfahrern bei Nacht im Vergleich zu Tagfahrten längere, vertikalere und insgesamt weniger Fixationen (Konstantopoulos et al., 2010). Werden diese Erkenntnisse auf den Bahnbereich übertragen, ist denkbar, dass die Erkennung von Gefahrenstellen, die in größerer horizontaler Entfernung vom Gleisbett liegen, in der Nacht schwerer fällt.

## 5.2 Hören/ auditive Wahrnehmung

Der Sinn des Hörens entsteht durch Schwingungen der Basalmembran in der Cochlea, welche durch Auftreffen einer Schallwelle auf die Membran entstehen. Vom Menschen werden Schallwellen zwischen 20 Hz und 20.000 Hz wahrgenommen. Hierbei benötigen Schallwellen tieferer Töne eine höhere Amplitude, um wahrgenommen zu werden, als hohe Töne. Frequenzunterschiede von 1/1000 Hz können bereits wahrgenommen werden (Becker-Carus und Wendt, 2017).

Die horizontale Lokalisation eines Tons erfolgt durch die Zeitdifferenz der Wahrnehmung zwischen beiden Ohren. Außerdem wird eine Frequenzveränderung des Tons wahrgenommen, die dadurch entsteht, dass die Schallwelle durch Auftreffen auf den Kopf verändert wird. Eine vertikale Lokalisierung geschieht durch die Form der Ohrmuschel. Durch sie entstehen verschiedene Frequenz-Dezibel-Funktionen, je nachdem, wo sich die Quelle einer Schallwelle befindet (Becker-Carus und Wendt, 2017).

Die auditive Wahrnehmung ist relevant für die Aufgaben der Tf insbesondere für die Kommunikation mit anderen Akteuren im Bahnbetrieb, z. B. mit Fahrdienstleitern, dem Disponenten des Eisenbahnverkehrsunternehmens oder mit anderen Tf. Zudem müssen Tf akustische Melder wahrnehmen können, z. B. den Sicherheitsfahrerschalter (Sifa) oder Meldungen der punktförmigen Zugbeeinflussung (PZB). Auch eine Überwachung des korrekten Funktionierens des Triebfahrzeugs erfolgt häufig durch ein auditives Beobachten der Umgebungsgeräusche. In den folgenden Abschnitten werden Einflussfaktoren auf die Leistungsfähigkeit der auditiven Wahrnehmung dargestellt.

Der Sinn des Hörens wird durch verschiedene Faktoren beeinflusst. Mit zunehmendem Alter nimmt die Fähigkeit, hochfrequente Töne zu hören, ab. Die absolute Hörschwelle, ab der ein Mensch in der Lage wäre, einen Ton wahrzunehmen, liegt bei  $2 \cdot 10^{-5}$  Pa (dies entspricht 0 dB; Mrowinski, Scholz und Stefens, 2017). Die Weber-Konstante beträgt für die Unterscheidung von Lautstärken  $k = 0,15$  und für die Unterscheidung von Tonhöhen bzw. Frequenzen  $k = 0,003$  (Becker-Carus und Wendt, 2017). Durch den anatomischen Aufbau der Ohrmuschel wird je nach Kopfposition und Frequenz der Schallwelle ein Band des Hörfrequenzspektrums abgedämpft (sog. ‚pinna notch‘). Außerdem geschieht beim Hören ähnlicher Frequenzen ein ‚Binding‘, also die Verbindung zu einem wahrgenommenen Wert. Dieses Binding wird unwahrscheinlicher, je weiter zwei Frequenzen voneinander entfernt sind. Maskierungseffekte entstehen dann, wenn gleichzeitig mit einem auditorischen Reiz oder kurz danach ein Reiz mit gleicher oder ähnlicher Frequenz präsentiert wird. Der eigentliche Reiz wird dann nicht wahrgenommen.

Menschen haben die Fähigkeit zur selektiven Wahrnehmung von auditorischen Reizen. Dies wird vor allem durch den Cocktailparty-Effekt (Cherry, 1953) demonstriert. Sind mehrere Geräuschquellen vorhanden, ist das menschliche Gehirn in der Lage, die relevanten akustischen Reize aus Störgeräuschen herauszufiltern. Dies ist insbesondere für die Aufgaben von Tf relevant. Tf müssen in der Lage sein, in einer Umgebung mit einem hohen Geräuschpegel wichtige auditive Informationen durch Melder und Zugfunk selektiv wahrzunehmen und zu priorisieren.

Weiterhin wird die auditive Wahrnehmung bzw. akustische Information nach Bubb (1977, zitiert nach Schlag und Heger, 2002) zur Einschätzung der eigenen Geschwindigkeit verwendet. Dem gegenüber steht die Einschätzung von Geschwindigkeitsdifferenzen, die mithilfe visueller Informationen erfolgt.

Auditive Wahrnehmung ist multimodal, d. h. sie kann durch gleichzeitig auftretende visuelle Reize beeinflusst werden (Becker-Carus und Wendt, 2017). Gerade in frühen Phasen der Wahrnehmung besteht eine Interaktion beider Modalitätsverarbeitungen. Durch die ergänzende Verarbeitung sowohl visueller als auch akustischer Informationen entsteht dabei ein Redundanzvorteil, was zu einer Verbesserung der Reaktionsleistung auf einen Stimulus führen kann (Becker-Carus und Wendt, 2017). Die Multimodalität der auditiven Wahrnehmung kann aber auch zu Täuschungen führen, wie beispielsweise durch den McGurk-Effekt demonstriert. Der McGurk-Effekt ist ein Phänomen, nachdem sich die Interpretation von Schallwellen je nach Vorliegen von Informationen über die Quelle der Schallwelle wie z. B. Mundbewegungen ändert (McGurk und MacDonald, 1976). So kann es sein, dass bei gleichzeitiger Präsentation einer Lautfolge ‚baba‘ und einer Lippenbewegung für ‚gaga‘ die Lautfolge als ‚dada‘ gehört wird.

## 5.3 Gleichgewicht

Der Gleichgewichtssinn des Menschen erfüllt mehrere Funktionen. Dazu gehören beispielsweise die Wahrnehmung und Berechnung der Körperhaltung des Menschen oder die Steuerung von Muskelreflexen (Brandes et al., 2019). Die Grundlage bildet das im Innenohr gelegene Vestibularorgan. Dieses besteht aus den Makulaorganen, die die Wahrnehmung linearer Geschwindigkeiten einschließlich Abbrems- und Beschleunigungsbewegungen ermöglichen, sowie den Bogengangorganen, die zur Wahrnehmung von Drehbewegungen dienen. Ergänzt wird der Gleichgewichtssinn durch den visuellen Sinn und die Propriozeption, die dabei helfen die Informationen des Vestibularorgans eindeutig zu interpretieren. Die Informationen aus den drei Quellen werden hauptsächlich im Hirnstamm und Kleinhirn verarbeitet (Brandes et al., 2019). Die Wahrnehmung von Abbrems-, Beschleunigungs- und Rotationsbewegungen ist als Grundlage der Geschwindigkeitswahrnehmung und Einschätzung des Zugverhaltens für Tf von großer Bedeutung. Das Vorliegen einer Anomalie des Vestibularorgans wird daher auch als Ausschlusskriterium für den Erhalt der Fahrberechtigung für Tf angegeben (Fünfte Verordnung zum Erlass und zur Änderung eisenbahnrechtlicher Vorschriften vom 29.04.2011; Bundesgesetzblatt I, Anlage 4, S. 721). (Becker-Carus und Wendt, 2017)

Der Kopf kann drei unterschiedliche Bewegungen ausführen: nicken, sich neigen und sich drehen. Diese Bewegungen werden um die drei Achsen des Raumes ausgeführt, wofür drei Bogengänge, welche in den drei Ebenen des Raums angeordnet sind, notwendig sind. Bei Drehbewegungen, welche lang andauern, entsteht aus einer anfänglichen Beschleunigung eine gleichförmige Drehgeschwindigkeit, welche über einen langen Zeitraum stabil beibehalten wird. Die gallertige Masse auf den Sinneshärchen der Bogengänge, die auch Cupula genannt wird, wird zunächst ausgelenkt, findet aber wegen ihrer Eigenelastizität während der konstanten Drehbewegung langsam wieder in die Ruhelage zurück. Bei einem schnellen Stopp der Bewegung, entsteht eine Auslenkung der Cupula in die entgegengesetzte Richtung, bevor diese nach relativ langer Zeit (10-30 s) wieder ihre Ruhelage erreicht (Klinke, 1995).

Zur Haltung des Gleichgewichts ist keine direkte Beteiligung des Bewusstseins nötig, da daran zwei Gruppen von Reflexen beteiligt sind. Die statischen Reflexe sorgen für spezifische Haltungen der einzelnen Glieder zueinander sowie bestimmte Körperstellungen im Raum. Statokinetische Reflexe sind Reaktionen auf Bewegungsreize und führen selbst wieder zu Bewegungen. Ein wichtiges Beispiel hierfür ist der vestibuläre Nystagmus, welcher eine visuelle Fixation von Gegenständen auch während Drehbewegungen ermöglicht. Dabei führt der Reflex die Augen langsam gegen eine Drehbewegung und ermöglicht eine schnelle Rückführung in Drehrichtung beim Anschlag der Augen. Zu beobachten ist dies bei einer Zugfahrt, bei der immer ein kurzer und schneller Nystagmus in Erscheinung tritt, bevor beim Vorbeifahren ein neues Objekt nach vorheriger langer Fixation fixiert werden kann (Klinke, 1995).

## 5.4 Haptische und taktile Wahrnehmung/Tastsinn

Der Tastsinn erlaubt die Wahrnehmung taktiler Reize wie beispielsweise Berührungen, Vibrationen oder Druck und lässt sich den Hautsinnen zuordnen (Hagendorf et al., 2011). Diese beinhalten neben dem Tastsinn noch die Temperatur- und die Schmerzwahrnehmung. Die den Hautsinnen zugrundeliegenden Informationen stammen dabei aus mehreren Rezeptionsquellen: den Mechanorezeptoren (zuständig für die Erkennung von Mustern/Formen, Druck, sowie Vibrationen), den kinästhetischen Rezeptoren (zuständig für die Propriozeption) und den Wärme- bzw. Kälterezeptoren.

Es wird weiterhin zwischen der eher passiven Propriozeption, die Auskunft über die Lage und Bewegung von Gliedmaßen des eigenen Körpers gibt, und der Haptik unterschieden. Als Haptik wird die aktive Wahrnehmung der Beschaffenheit von Objekten verstanden (Hagendorf et al., 2011). Neben den genannten Rezeptoren ist dabei auch die Motorik zur Steuerung der tastenden Hände involviert und die Informationen werden in kognitiven Prozessen integriert und interpretiert, um Rückschlüsse auf die Beschaffenheit des Objekts zu ermöglichen (Hagendorf et al., 2011).

Im Rahmen der Aufgaben der Tf sind unter anderem die Informationen relevant, die von den Mechanorezeptoren wahrgenommen werden. Erfahrene Tf können durch die Wahrnehmung der Vibrationen des Triebfahrzeugs Informationen über die Betriebssituation gewinnen, z. B. über die aktuelle Haftreibung.

Die Wahrnehmungsschwelle und simultane Raumschwelle, also der kleinste Abstand zweier Punkte, die noch als solche wahrgenommen werden können, sind abhängig von der Region des Körpers. Die empfindlichsten Körperteile sind aufgrund hoher Dichte an Mechanorezeptoren die Lippen, die Zunge und die Fingerspitzen. Laut Becker-Carus und Wendt (2017) gilt als Wahrnehmungsschwelle zur Unterscheidung des Hautdrucks die Weber-Konstante  $k = 0,14$ . Die Haut kann Infrarotwellen im Wellenlängenbereich von  $10^{-4}$  bis  $10^{-3}$  m wahrnehmen. Gegenüber Druckvibrationen, am stärksten bei Vibrationsfrequenzen zwischen 150 Hz und 300 Hz, ist die Haut besonders empfindlich. Die Unterschiedsschwelle für die Wahrnehmung von Änderungen der Vibrationsfrequenz steigt bei Frequenzen über 100 Hz rapide an und ist im Bereich niedriger Frequenzen am besten.

Bei Bewegungen werden viele Hautrezeptoren aktiviert, weswegen eine tastende Hand eine bessere Wahrnehmung hat als eine ruhende. Ein weiterer Grund hierfür ist, dass die Adaptation der Rezeptoren weitgehend vermindert wird, wodurch detailliertere Informationen über die Umgebung aufgenommen werden können (Gibson, 1962). Fremdverursachter, mechanischer Druck wird als stärker wahrgenommen als selbstverursachter derselben Stärke. Bei länger andauernder, gleichartiger Reizeinwirkung ist die Adaptation des taktilen Wahrnehmungssystems groß, solange keine Bewegung an dieser Stelle erfolgt, wie z. B. bei einem Ring am Finger. Eine schnelle und vollständige Adaptation ist außerdem bei mäßigen Temperaturunterschieden zu beobachten, sodass nach wenigen Minuten keine Wärme oder

Kälte aktiv wahrgenommen wird. Schmerzempfindungen involvieren viele personenbezogene Erfahrungen und werden durch Gewohnheiten, Erwartungen, frühere Erfahrungen und der Kulturzugehörigkeit einer Person mitbestimmt (Becker-Carus und Wendt, 2017).

## 5.5 Zusammenfassung

In den vorhergehenden Abschnitten wurde für jeden Sinn herausgearbeitet, welche Einflüsse die Leistungsfähigkeit dieses Sinns bei der Wahrnehmung beeinflussen. Dabei können mehrere Sinne von einem Einflussfaktor betroffen sein. Tabelle 2 enthält eine Zusammenstellung der Sinne, jeweiligen Einflussgrößen und des Effekts des Einflusses. Weitere Einflussfaktoren, die nicht speziell den Sinnen zugeordnet wurden, sind im Kapitel 17 zu finden.

TABELLE 2: IDENTIFIZIERTE EINFLUSSFAKTOREN FÜR DIE MENSCHLICHEN SINNE

<b>Sehen/Visuelle Wahrnehmung</b>	<b>Hören/auditive Wahrnehmung</b>	<b>Gleichgewichtssinn</b>	<b>Haptische und taktile Wahrnehmung</b>
<b>Gesichtsfeld und Tunnelblick</b>	Absolute Hörschwelle	Vestibularorgan im Innenohr	Wahrnehmungsschwelle
<b>Hell-Dunkel-Adaptation und Blendung</b>	Alter	Visueller Sinn und Propriozeption	Frequenz der Vibrationen
<b>Kontraste und Größe</b>	Anatomischer Aufbau des Außenohrs	Bewegung des Kopfes	Bewegung und Adaptation
<b>Entfernungswahrnehmung</b>	Maskierungseffekte	Statische und Statokinetische Reflexe	Infrarotwellen
<b>Geschwindigkeit</b>	Cocktailparty-Effekt (Selektive Wahrnehmung)	Vestibulärer Nystagmus	Fremd- und Selbstverursachter Druck
<b>Alter</b>	Geschwindigkeitswahrnehmung		Adaptation bei langanhaltenden Reizen
<b>Belastung</b>	Multimodalität und Täuschung		Schmerzempfindungen
<b>Ermüdung</b>	Binding ähnlicher Frequenzen		Körperregionen
<b>Fahrererfahrung</b>			
<b>Exzentrizität</b>			
<b>Bewegung und Sichtfeld</b>			
<b>Obstruktion</b>			
<b>Beleuchtung</b>			

## 5.6 Anforderungen an die Wahrnehmungsleistung von Tf

In der fünften Verordnung zum Erlass und zur Änderung eisenbahnrechtlicher Vorschriften vom 29.04.2011 (Bundesgesetzblatt I, Anlage 4, S. 721) ist festgelegt, dass eine Fern-Sehschärfe mit oder ohne Sehhilfe von 1,0 und mindestens 0,5 für das schlechtere Auge vorliegen darf. Eine Korrekturlinse darf eine maximale Korrektur von +5 bis -8 Dioptrien aufweisen. Das Sehvermögen auf kurze und mittlere Entfernungen muss ausreichend sein. Die Farbwahrnehmung muss normal sein und das Gesichtsfeld vollständig. Sowohl das Sehvermögen beider Augen, als auch das binokulare Sehvermögen muss effektiv sein. Die Prüfung der Erkennung farbiger Signale erfolgt auf der Grundlage einzelner Farben, nicht auf der Grundlage relativer Unterschiede. Die Kontrastempfindlichkeit muss gut sein und es dürfen keine fortschreitenden Augenkrankheiten vorliegen. Linsenimplantate, Keratotomien und Keratektomien sind nur zulässig, wenn sie jährlich oder in vom Arzt festgelegten regelmäßigen Abständen überprüft werden. Es darf keine Überempfindlichkeit gegen Blendung vorliegen und farbige Kontaktlinsen und fotochromatische Linsen sind nicht zulässig.

Zudem ist aufgeführt, dass genug Hörvermögen für ein Telefongespräch und die Fähigkeit, akustische Warnsignale und Funkmeldungen zu hören, vorliegen muss. Außerdem darf kein Hördefizit von über 40 dB bei 500 Hz und 1.000 Hz vorliegen. Auf dem Ohr, welches die schlechtere Schallleitung aufweist, darf kein Hördefizit von über 45 dB bei 2.000 Hz vorliegen. Es wird außerdem jeder mit einer Anomalie des Vestibularorgans ausgeschlossen. Da es als Tf notwendig ist, Mitteilungen laut und deutlich auszutauschen, darf ebenso keine chronische Sprachstörung bestehen. Die Verwendung von Hörhilfen ist in bestimmten Fällen zulässig.

## 6 Ableitung von Szenarien

### 6.1 Auswahl des zu untersuchenden Sinns

In den Anforderungen an die Wahrnehmungsleistung von Tf für die Erteilung eines Triebfahrzeugführerscheins (Fünfte Verordnung zum Erlass und zur Änderung eisenbahnrechtlicher Vorschriften vom 29.04.2011; Bundesgesetzblatt I, Anlage 4, S. 721) werden die Sinne Sehen und Hören berücksichtigt, wobei die Anforderungen an den Sinn des Sehens überwiegen. Dies spiegelt sich auch in der in diesem Projekt erstellten Beschreibung der Tf-Aufgaben wider (siehe Kapitel 3). Insgesamt stehen rund 69 % der ermittelten Aufgaben von Tf im Zusammenhang mit der visuellen Wahrnehmung. Auch (Brandenburger et al., 2017) führt aus, dass eine der Hauptaufgaben der Tf die Informationsaufnahme aus den unterschiedlichen visuellen Reizen in der Arbeitsumgebung ist. Aus diesem Grund wurde entschieden, sich in der weiteren Untersuchung auf den Sinn Sehen zu konzentrieren.

### 6.2 Auswahl einer Standardaufgabe

Im Rahmen der im Projekt geplanten Simulatoruntersuchung war es aus zeitlich-organisatorischen Gründen nicht möglich, alle mit dem Sehsinn im Zusammenhang stehenden Standardaufgaben der Tf (siehe Tabelle 1) zu untersuchen. Das Projekt führte die Analyse prototypisch für eine Standardaufgabe durch, eine Übertragbarkeit der grundsätzlichen Überlegungen scheint gegeben. Da jede Standardaufgabe eine Aggregation verschiedener Einzelaufgaben von Tf darstellt, wurde bereits durch die Untersuchung nur einer Standardaufgabe der Blick auf einen sinnlogischen Teil der Tf-Aufgaben in ihrer Gesamtheit möglich. Die Auswahl der zu untersuchenden Standardaufgabe wurde mithilfe einer mehrstufigen Filtermethode durchgeführt.

In der ersten Stufe wurden Ergebnisse aus einer Studie zu möglichen Automatisierungsaspekten und die Funktionsteilung Mensch-Technik im hochautomatisierten Schienenverkehr herangezogen (Brandenburger et al., 2017). Aufbauend auf diesen Ergebnissen wurde beurteilt, welche der Standardaufgaben im hochautomatisierten Schienenverkehr mit einer hohen Wahrscheinlichkeit automatisiert werden und welche der Bearbeitung durch den Menschen überlassen bleiben (Automatisierbarkeit). Diese Einschätzung der Standardaufgaben wurde vorgenommen, da im Projekt die Leistungsfähigkeit menschlicher Sinne bei Aufgaben beurteilt werden soll, die zukünftig von Sensoren übernommen werden sollen. Eine Bearbeitung von Standardaufgaben, die nicht automatisiert werden, ist demnach inhaltlich nicht zielführend (Tabelle 3). In der ersten Stufe der Filterung wurden die Standardaufgaben *Fahren/ Geschwindigkeit regeln*, *Türschluss*, *Monitoring Fahrzeug* und *Monitoring Infrastruktur* ausgewählt. Die Standardaufgaben *Diagnose* und *Kuppeln* sind eingeschränkt automatisierbar und wurden nicht weiter betrachtet.

Die nächste Stufe der Filtermethode basierte auf der Umsetzbarkeit im Fahrsimulator. Die Fähigkeiten der Fahrsimulatoren des DLR und des BBI wurden identifiziert und die Standardaufgaben daraufhin überprüft, ob sie in den Simulatoren umgesetzt werden können. In dieser Filterstufe entfiel die Standardaufgabe *Türschluss*, da diese in keinem der Simulatoren sinnvoll umgesetzt werden kann.

Nach diesen Filterstufen verblieben die Standardaufgaben *Fahren/ Geschwindigkeit regeln*, *Monitoring Fahrzeug* und *Monitoring Infrastruktur*. Die weitere Auswahl wurde basierend auf der Analyse des Risikopotentials getroffen. Für die Aufgaben der Tf wurde eine qualitative Unfallanalyse durchgeführt.

In Anlehnung an die VDE V 0831-103 wurden die Auswirkungen beim Versagen, d. h. Unfallart und Schadensausmaß, ermittelt (DIN VDE V 0831-103:2020). Bei den meisten der 28 betrachteten Aufgaben standen die Unfallursachen im Zusammenhang mit den Überwachungsaufgaben der Tf. Etwa die Hälfte der Aufgaben mit Risikopotential ließ sich als Teil der Standardaufgabe *Monitoring Infrastruktur* und die andere Hälfte als Teil der Standardaufgabe *Monitoring Fahrzeug* einstufen (Kapitel 17, Anhang 2). Aus diesem Grund sollte eine dieser beiden Standardaufgaben für die weitere Bearbeitung ausgewählt werden.

Als Basis für die Entscheidung zwischen den Standardaufgaben *Monitoring Infrastruktur* und *Monitoring Fahrzeug* wurden Unfalluntersuchungsberichte der BEU (BEU, o. J.) analysiert, um die Rolle der Tf bei Unfällen von 2014 bis 2020 zu identifizieren. Laut den Untersuchungsberichten zu Personenumfällen und Unfällen an Bahnübergängen wurde bei zwei von 19 Unfällen Tf als verantwortlich ermittelt. Zwei weitere Unfälle waren auf eine fehlerhafte Signalüberwachung und eine fehlerhafte Abfahrtsbereitschaft zurückzuführen. Von 24 Entgleisungsunfällen ereignete sich einer aufgrund der mangelnden Signalbeobachtung durch den Tf. Zusätzlich wurden in einigen Berichten die Lichtverhältnisse, Nebel und Sichteinschränkungen als beitragende Faktoren genannt. Bei 33 Unfällen mit Zugkollisionen wurden bei 12 Vorfällen ein oder mehrere Fehler der Tf festgestellt. Signalbeobachtung und unzulässiges Vorbeifahren an haltzeigenden Signalen waren die häufigsten Fehler der Tf. Darüber hinaus standen vier Vorfälle teilweise oder direkt mit der manuellen Handhabung in Zusammenhang, z. B. eine fehlerhaft durchgeführte Bremsprobe.

Zusammenfassend kann festgehalten werden, dass die Mehrheit der betrachteten Ereignisse auf Fehler bei der Überwachung der externen Infrastruktur zurückzuführen ist. Aus diesem Grund wurde abschließend die Standardaufgabe *Monitoring Infrastruktur* für die weitere Untersuchung in der Simulatorstudie ausgewählt. Die Operationalisierung sah vor, dass Tf eine Strecke befahren, die Infrastruktur beobachten und auf Objekte reagieren. Dies entspricht den Aufgaben Streckenbeobachtung aus Tabelle 1, beschrieben in der Fahrdienstvorschrift Abschnitt 408.2341 (Richtlinie 408.21-27:2014)

TABELLE 3: STANDARDAUFGABEN UND DEREN JEWEILIGE AUTOMATISIERBARKEIT UND UMSETZBARKEIT IM SIMULATOR

<b>Standardaufgabe</b>	<b>Automatisierbarkeit</b>	<b>Umsetzbarkeit im Simulator</b>
<b>Fahren/V Regeln</b>	ja	ja
<b>Diagnose</b>	teilweise	nein
<b>Kommunikation Reisende</b>	nein	Ja
<b>Kommunikation betrieblich</b>	nein	nein
<b>Kuppeln</b>	vielleicht	nein
<b>Instandsetzung</b>	nein	nein
<b>Evakuierung</b>	nein	nein
<b>Türschluss</b>	ja	nein
<b>Monitoring Fahrzeug</b>	ja	ja
<b>Monitoring Infrastruktur</b>	ja	ja

## 6.3 Unabhängige Variablen für die Simulatorstudie

Die Qualität der Aufgabenausführung wird entscheidend durch die Wirkung von Einflussfaktoren beeinflusst. Die Simulatorexperimente waren so aufzusetzen, dass der Einfluss der wichtigsten PSF abgebildet wird und die Bandbreite der menschlichen Leistung, gemessen als Reaktionszeit, bei Anwendung des Sinnes Sehen erfasst werden konnte. Die relevantesten Einflussfaktoren wurden als unabhängige Variablen ausgewählt, d. h. im Rahmen der Simulatorstudie variiert. Die Variation der Variablen ermöglichte es, Zusammenhänge zwischen diesen Faktoren und der Reaktionszeit zu untersuchen.

Die basierend auf die Literaturrecherche identifizierten PSF wurden durch das Konsortiums hinsichtlich der erwarteten Relevanz bewertet (Kapitel 17, Anhang 3). Für die Analyse wurde ein Likert-Skalen-basiertes Benchmarking verwendet. Für jeden Faktor wurde die Wichtigkeit des Faktors von „gar nicht wichtig“ bis „sehr wichtig“ bewertet.

Ergänzend zur Relevanzbewertung der PSF wurden die Ergebnisse der durchgeführten Literaturrecherche zum Sehsinn herangezogen (Tabelle 2). Einzelne Einflussfaktoren auf die Leistung der visuellen Wahrnehmung wurden im Abschnitt 5.1 beschrieben. Die identifizierten Faktoren sind in Tabelle 4 aufgeführt. Dabei handelt es sich um die Eigenschaften der Stimuli oder um interne oder externe Faktoren, die die Wahrnehmungsleistung beeinflussen.

TABELLE 4: EIGENSCHAFTEN DES WAHRZUNEHMENDEN REIZES UND FAKTOREN, DIE DIE WAHRNEHMUNG BEINFLUSSEN

Kategorie	Einflussfaktoren	Beispielmerkmale
<b>Eigenschaften des wahrzunehmenden Reizes</b>	Größe und Form des Reizes	klein
		groß
	Einsehbarkeit des Reizes	Einschränkung durch Kurven
		Einschränkung durch Bewuchs
	Position des Reizes (Exzentrizität)	nah neben dem Gleis
		auf dem Gleis
		fern vom Gleis
	Kontrast	guter Kontrast
schlechter Kontrast		
<b>Interne und externe Faktoren, die die Wahrnehmung beeinflussen</b>	Variation der Aufmerksamkeit/des visuellen Suchverhaltens	Fahren auf Sicht vs. Fahren mit Movement Authority
		Bahnübergänge vs. freie Strecke
	Blickverteilung	Zugsicherungssysteme PZB vs. LZB/ETCS
	Einengung des Blickfelds	Geschwindigkeit
		reizarme Umgebung
		Fahren durch Schallschutzwände vs. freie Umgebung
	Anspannung/Stress/Zeitdruck	betriebliche Störungen
Müdigkeit	Fahrtdauer	
Ablenkung	sekundäre Aufgabe	

Im Ergebnis der Analyse wurden die in Tabelle 5 dargestellten Einflüsse als unabhängige Variablen ausgewählt. Die betrieblichen Parameter sowie die physikalischen Eigenschaften des Reizes wurden im Rahmen der Simulatorexperimente aktiv variiert. Die persönlichen Einflussfaktoren wurden im Gegensatz dazu nicht aktiv variiert, sondern lediglich miterhoben. So konnten ihre möglichen Auswirkungen auf die Leistung berücksichtigt werden. In den folgenden Abschnitten werden die Variationsstufen der unabhängigen Variablen genauer beschrieben.

TABELLE 5: UNABHÄNGIGE VARIABLEN

Kategorie	Unabhängige Variable	Variationsstufen
<b>Persönliche Einflussfaktoren</b>	Alter	Erhebung ohne aktive Variation
	Müdigkeit	
	Arbeitserfahrung	
<b>Betriebliche Parameter</b>	Geschwindigkeit	40 km/h
		100 km/h
		160 km/h
	Zugsicherungssystem	PZB
		ETCS
		Fahren auf Sicht
<b>Physikalische Eigenschaften des Reizes</b>	Größe	klein
		groß
	Kontrast	schlecht
		gut

### 6.3.1 Persönliche Einflussfaktoren

Es wurden die folgenden Daten von den Teilnehmenden abgefragt: Alter, Arbeitserfahrung (Berufserfahrung sowie die Vorerfahrung mit PZB und ETCS). Diese Variablen wurden nicht aktiv variiert. Im Kapitel 5 wurde beschrieben, dass die Fähigkeiten der visuellen Wahrnehmung durch physiologische und optische Veränderungen des Auges mit dem Alter abnehmen. Die in den Experimenten erhobenen demografischen Daten umfassen daher das Alter der Teilnehmenden. Das Ausmaß der Einschränkung der Sehfähigkeiten ist jedoch interindividuell unterschiedlich. Die Arbeitserfahrung kann durch verbesserte Antizipation einen Einfluss auf die Leistungsfähigkeit haben.

Zusätzlich wurde die Ermüdung der Teilnehmenden vor Beginn der Fahrt und nach jedem Versuchsblock erhoben. Müdigkeit wurde in den Experimenten nicht aktiv variiert, jedoch wurde die Veränderung der Müdigkeit durch eine psychomotorische Wachsamkeitsaufgabe (Karolinska-Sleepiness Scale) gemessen (Shahid et al., 2012). Die subjektive mentale Beanspruchung wurde nach jeder Fahrt mithilfe der NASA-TLX-Skala gemessen (Hart & Staveland, 1988). Nach Desmond, Hancock und Monette (1998) können längere Arbeitsphasen mit hoher Arbeitsbelastung zu aktiver Ermüdung führen und Situationen mit kontinuierlich niedriger Arbeitsbelastung können zu passiver Ermüdung führen. Ermüdung kann bei Tf als Folge von Arbeitszeit und -bedingungen (DaCoTA, 2012), Schlafverlust, (Dorrian et al., 2007), geringer Aufgabenbelastung und Monotonie (Stein und Naumann, 2016) auftreten. Die Auswirkungen von Müdigkeit auf die sensorische Leistung wurden in Kapitel 5 erläutert. In Kapitel 5.1.2.2 wurde erläu-

tert, dass sich auch die Beanspruchung, die durch eine Nebenaufgabe oder Situationen mit hoher Informationsdichte entsteht, negativ auf die Gefahrenerkennung auswirken kann. Ergänzend wurde auch ein psychomotorischer Wachsamkeitstest (PVT) durchgeführt.

### 6.3.2 Betriebliche Parameter

Die Aufmerksamkeitsverteilung zwischen Führerstand und Umfeld steht in direktem Zusammenhang mit dem Risiko, streckenseitige Gefahrenquellen zu übersehen, die nicht vom Signalsystem selbst überwacht oder gesteuert werden. In der Literatur finden sich Studien, die die Auswirkungen von Zugsicherungs- und Fahrassistenzsystemen auf die Aufmerksamkeitsallokation untersuchen (Marinakos, Sheridan und Multer, 2005, Hely et al., 2015, Brandenburger, Naumann und Jipp, 2019, Naghiyev et al., 2014). Im Projekt wurden zwei Zugsicherungssysteme mit vermuteter unterschiedlicher Aufmerksamkeitsallokation eingesetzt, um deren Auswirkung auf die Wahrnehmungsleistung zu untersuchen. Diese Systeme sind die punktförmige Zugbeeinflussung (PZB), bei der der Fahrer weiter mit Außensignalisierung fährt und durch die PZB lediglich die korrekte Reaktion bei einschränkenden Signalbegriffen prüft und ggf. den Zug zum Halten bringt, und ETCS Level 2 mit Führerstandssignalisierung ohne Außensignale, bei dem weitgehend alle Informationen im Führerstand angezeigt werden.

Der Zusammenhang zwischen Geschwindigkeit, visueller Orientierung und dem Blickfeld wurde in Kapitel 5 beschrieben. Beim Fahren mit sehr hohen Geschwindigkeiten erreicht die Menge der visuellen Informationen die Verarbeitungsgrenze des Menschen (Guo et al., 2015). Auch wurde festgestellt, dass mit steigender Geschwindigkeit die Vertikalität der Augenbewegungen zunimmt und die Horizontalität der Augenbewegungen abnimmt (Suzuki et al., 2019). Um den Einfluss der Geschwindigkeit auf die Sehleistung von Tf zu bewerten, wurden drei unterschiedliche Höchstgeschwindigkeiten für die simulierten Streckenabschnitte gewählt, nämlich 40 km/h, 100 km/h und 160 km/h. Die streckenseitige Signalisierung kann bis zu einer Geschwindigkeit von 160 km/h verwendet werden, bei höheren Geschwindigkeiten ist Führerstandssignalisierung zwingend notwendig. Mit den beiden gewählten Geschwindigkeitsstufen 100 km/h und 160 km/h wird der typische Anwendungsbereich im Fern- und Regionalverkehr abgedeckt.

Es gibt betriebliche Situationen (z. B. Räumungsprüfung oder unklarer Zustand der Infrastruktur), in denen das Fahren auf Sicht (FaS) erforderlich ist. Je nach Sichtverhältnissen wird die Geschwindigkeit durch den Tf selbstständig zwischen Schrittgeschwindigkeit und maximal 40 km/h geregelt. Zur Bewertung der Wahrnehmungsleistung beim Fahren auf Sicht wurde ein zusätzliches Szenario berücksichtigt.

Die Geschwindigkeit 100 km/h wurde sowohl auf der ETCS- als auch auf der PZB-Strecke umgesetzt. Die Geschwindigkeit 40 km/h wurde sowohl für die PZB-Strecke als auch beim FaS umgesetzt. Nur die ETCS-Strecke erlaubte das Fahren mit 160 km/h. Die wahrzunehmenden Würfel (s. u.) auf den ETCS- und PZB-Strecken waren so platziert, dass die Geschwindigkeit zum Zeitpunkt der Reaktion in zwei verschiedenen Geschwindigkeitsklassen lag.

### 6.3.3 Physikalische Eigenschaften des Reizes

Tf müssen im Rahmen der Aufgabenausführung alle Objekte, die den Lichtraum des Fahrzeugs verletzen bzw. verletzen können, erkennen und ggf. entsprechende Reaktionen ableiten. Aufgrund der unbegrenzten Vielzahl an Objekten ist es nicht möglich, diese alle im Experiment zu berücksichtigen. Der Schutz von Personen ist ein besonders hohes Ziel des Bahnbetriebs. Daher muss der Tf vor allem alle Menschen im oder beim Lichtraum erkennen. Mit wachsender Geschwindigkeit wird für Passanten

wegen der Aerodynamik auch der Gefahrenraum größer. Im Konsortium wurde unter Berücksichtigung der bei den Konsortiumsmitgliedern gültigen ethischen Regeln abgestimmt, dass menschliche Figuren als Objekte zur Erkennung nicht herangezogen werden, um eine potentielle Traumatisierung der Tf zu vermeiden. Es wurde entschieden, dass als Objekte Würfel verwendet werden.

Die Größe des wahrzunehmenden Reizes beeinflusst die mögliche Leistung bei einer Wahrnehmung des Reizes (siehe Abschnitt 5.1.1.3) variiert. Die Auswahl der Objektgrößen orientierte sich an den Menschen- und Großtiergrößen basierend auf Klassifikationen des Projekts. Um statistisch aussagekräftige Ergebnisse zu ermöglichen, wurde die Zahl der im Versuch verwendeten Kombinationen begrenzt. Es wurden in den Experimenten Reize mit zwei unterschiedlichen Größen verwendet. Die Größe der Stimuli wurde über die Kantenlänge der Würfel (90 cm als „klein“ vs. 180 cm als „groß“) variiert. Die Auswahl der Objektgrößen orientierte sich an den Menschen- und Großtiergrößen basierend auf Klassifikationen des Projekts ATO-Sensorik (Leinhos, Flatt und Witte, 2022).

Wie in Kapitel 5.1.1.3 erläutert, ist der Kontrast einer der Parameter, die die Wahrnehmung eines Reizes beeinflussen. Der Kontrast der Würfel zum Hintergrund wurde durch die Farben der Würfel variiert. Es wurden eine kontrastreiche und eine kontrastarme Bedingung umgesetzt. Für die kontrastreiche Bedingung wurde ein helles Orange gewählt, das auf der Farbe von Sicherheitswesten basiert (HEXCode #f18e2a). Für die Ermittlung der kontrastarmen Farbe wurden zunächst der Leuchtdichtekontrast und der Farbkontrast zwischen der kontrastreichen Stimulusfarbe und einer durchschnittlichen Hintergrundfarbe berechnet. Für die kontrastarme Farbe wurde dann ein Farbton gewählt, der etwa der Hälfte des Leuchtdichte- und Farbkontrasts der kontrastreichen Farbe entsprach, ein orange-brauner Ton (HEX-Code #9d6830). Die Hintergrundfarbe wurde anhand der unmittelbaren Umgebungsfarbe jedes Objekts berechnet. Der Farbkontrast wurde anhand der von Mokrzycki und Tatol (2011) vorgeschlagenen Delta-E-Formel im RGB-Farbraum berechnet. Obwohl es andere Farbunterschiedsformeln gibt, die auf verschiedenen Farbräumen wie CIELAB basieren, wurde die folgende Formel aufgrund ihrer Effizienz in der Berechnung und des Anspruchs an Präzision im Projekt ausgewählt.

$$\Delta_E = \sqrt{\left(2 + \frac{R_{Sr}}{256}\right) \Delta R^2 + 4\Delta G^2 + \left(2 + \frac{255 - R_{Sr}}{256}\right) \Delta B^2}$$

$\Delta_E$ : Farbunterschied

$\Delta R, \Delta G, \Delta B$ : Abstand zwischen Punkten im Raum, der die RGB-Farben repräsentiert

$R_{Sr}$ : Durchschnittswert der R-Komponenten in den verglichenen Farben

Die Tabelle 6 zeigt die Kontrastberechnung für zwei Stimulieigenschaften als Beispiel. Für den Stimulus A gibt es große Objekte mit hohem Kontrast (Wh) und niedrigem Kontrast (Wn). U1 und U2 bezeichnen die durchschnittlichen RGB-Werte des Würfelhintergrunds. Die Delta-E-Werte werden berechnet, um den Farbunterschied zwischen dem Würfel und seiner Umgebung zu ermitteln. Das Verhältnis zwischen den kontrastreichen und den kontrastarmen Bedingungen ist in der letzten Zeile angegeben. Dieser Wert liegt im Bereich von 45 % bis 51 % für alle Würfel. Die Abbildung 4 zeigt zwei Stimuli mit hohem und niedrigem Kontrast als Beispiel.

TABELLE 6: BERECHNUNG DES KONTRASTVERHÄLTNISSES VON ZWEI VERSCHIEDENEN BEDINGUNGEN (BBI)

$\Delta_{Ei}$ (Würfel-Umgebung)		Bedingung A (groß)	Bedingung B (klein)
Wh-U1	$\Delta_{E1}$	217	233
Wn-U2	$\Delta_{E2}$	97	112
Kontrastverhältnis	$\Delta_{E2}/\Delta_{E1}$	0,45 (45 %)	0,48 (48 %)

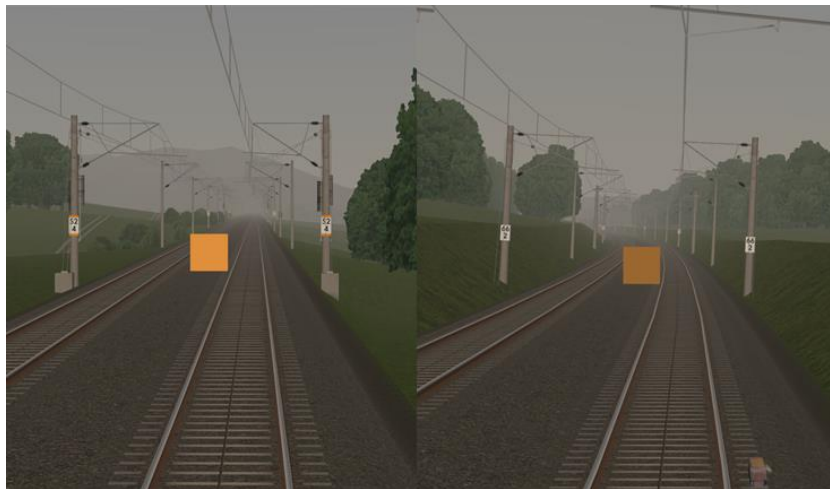


Abbildung 4: Kontrastvarianten der Würfel der BBI-Simulatorstudie

Der Kontrast der Farben wurde zusätzlich im DLR-Simulator über eine Messung der Beleuchtungsstärke mittels Luxmeter auf der Beamer-Leinwand kontrolliert (siehe Tabelle 7). Dafür wurde die Beleuchtungsstärke für den Würfel sowie für den umgebenden Bereich (links, rechts, oberhalb und unterhalb des Würfels), das Gleis und den Himmel exemplarisch an zwei Orten auf der Versuchsstrecke bestimmt (Abbildung 5). Zum Vergleich des Kontrasts beider Kontrastbedingungen kann die Beleuchtungsstärke des Würfels  $W$  geteilt durch den Mittelwert der Beleuchtungsstärke der Würfelumgebung  $M_U$  betrachtet werden. Die Differenz dieser Werte für die kontrastreiche und kontrastarme Bedingung beträgt  $2,67 - 1,75 = 0,92$  an Ort 1 und  $3,09 - 2,15 = 0,94$  an Ort 2. Somit ist der Kontrastunterschied der Würfel im Vergleich zur Umgebung zwischen den beiden Größenbedingungen an beiden Orten nahezu identisch.

TABELLE 7. BELEUCHTUNGSSTÄRKEN DER WÜRFEL UND UMGEBUNG DER DLR-SIMULATORSTUDIE

Gemessener Bereich (in Lux)	Ort 1 (Würfel groß)		Ort 2 (Würfel klein)	
	kontrastreich	kontrastarm	kontrastreich	kontrastarm
Würfel (W)	138	87	160	96
links	55	45	36	34
rechts	70	60	60	45
oben	32	46	55	45
unten	50	48	56	55
Mittelwert Umgebung ( $M_U$ )	51,75	49,75	51,75	44,75
Himmel	284	290	260	290
Gleis (G)	54	50	55	56
$W/M_U$	2,67	1,75	3,09	2,15
$W/G$	2,56	1,74	2,91	1,71

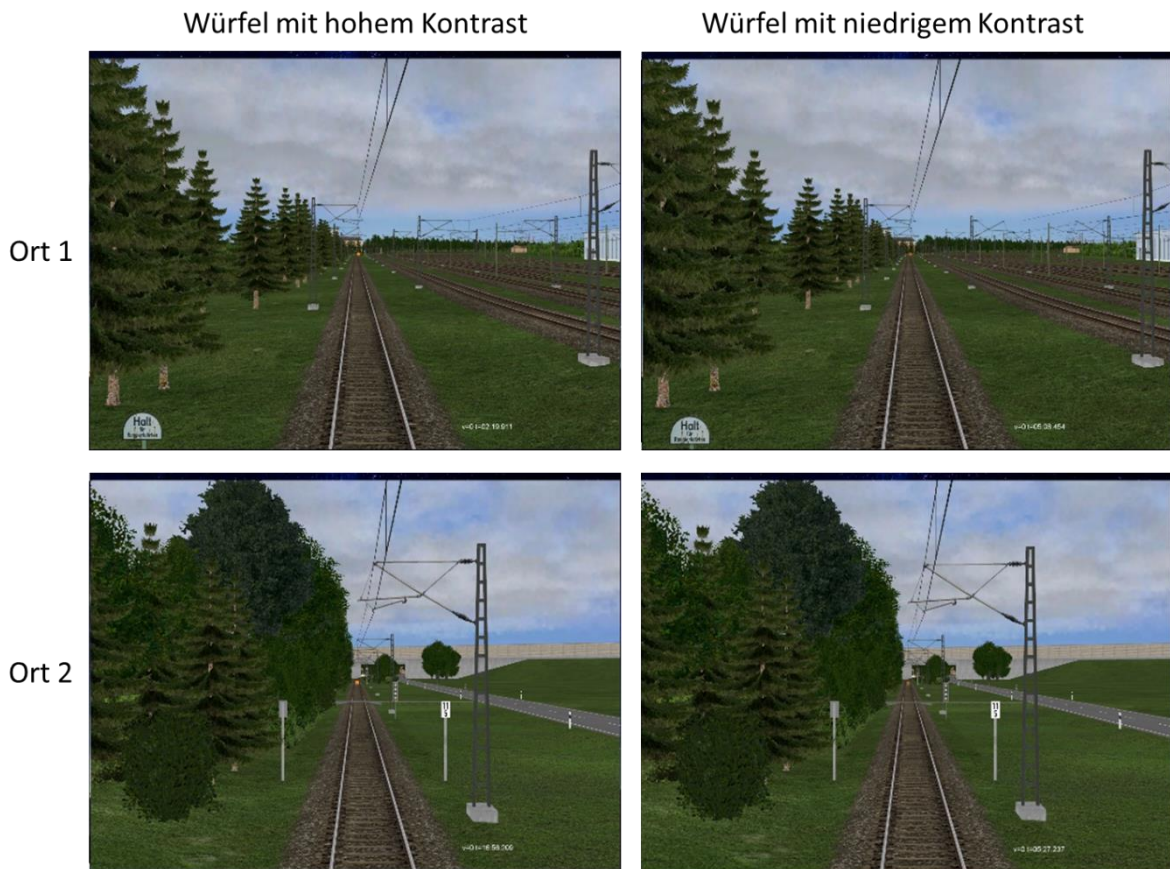


Abbildung 5: Größen- und Kontrastvarianten der Würfel der DLR-Simulatorstudie.

Um andere Variablen wie Streckengeometrie und Umgebung zu kontrollieren, wurden freie Streckenabschnitte mit geringer oder keiner Steigung als Standorte für die Würfel ausgewählt. Es gab keine Würfel in Tunneln oder an den Bahnhöfen.

In der BBI-Studie wurden die Stimuli rechts oder links von den Gleisen platziert, mit einem maximalen Abstand von 3 m von der Gleismitte. Der Stimuli befand sich damit weiterhin im zentralen Sichtfeld der Tf. In der DLR-Studie wurden die Stimuli in der Mitte der Gleise platziert. Befand sich der Zug in einer Entfernung von 800 m von der vorgesehenen Position des Stimulus, wurde dieser eingeblendet, sodass sichergestellt war, dass jeder Stimulus mit einem gleichen Abstand erstmals gesehen werden konnte. Die Entscheidung für einen Abstand von 800 m basierte darauf, dass der Stimulus, empirisch getestet, ab 800 m bei Bekanntsein des Auftauchorts gerade so sichtbar war. Es wurde darauf geachtet, dass der Stimulus vom Moment seines Erscheinens bis zum Passieren der Stimulusposition sichtbar war, damit die Teilnehmenden die Möglichkeit hatten, den Stimulus zu jedem Zeitpunkt der Annäherung zu sehen und darauf zu reagieren. Der Stimulus war bei seinem ersten Erscheinen noch so weit entfernt von der aktuellen Zugposition, dass Pop-Out-Effekte durch das Erscheinen des Stimulus vermieden wurden. Um zu vermeiden, dass durch eine vermeintliche Kollision mit dem Stimulus Erinnerungen an Unfälle mit Gegenständen oder Personen auf dem Gleis ausgelöst werden, sprang der Stimulus beim DLR-Simulator vor der Ankunft des Zuges am Stimulusort an eine Stelle neben dem befahrenen Gleis. Um weitgehend realistische Bedingungen zu schaffen, wurde darauf geachtet, dass die Abstände zwischen den Würfeln möglichst unregelmäßig, mit einem Mindestabstand von ca. 2,5 Minuten, eingehalten wurden.

## 6.4 Abhängige Variablen für die Simulatorstudie

Als Messgröße, also die abhängige Variable, wurde die Reaktionszeit auf einen Reiz, d. h. im Projekt die Reaktionszeit für die visuelle Wahrnehmung des Objekts, gewählt. Der Begriff Reaktionszeit wird häufig in der experimentellen Psychologie und der Informationsverarbeitungsforschung verwendet. Boff und Lincoln (1988) definieren die Reaktionszeit als die Zeit vom Einsetzen eines Reizes bis zum Beginn der Reaktion auf den Reiz durch einen einfachen motorischen Akt (z. B. einen Knopfdruck). Wichtig zu berücksichtigen ist, dass „Reaction time“ und „Response time“ in der englischsprachigen Fachliteratur manchmal synonym verwendet werden: Nur wenn auch die Zeit für die Handlung enthalten ist, sollte die Gesamtdauer als "Response Time" bezeichnet werden. Die Reaktionszeit oder response time umfasst also die Zeit für die mentale Verarbeitung des wahrgenommenen Reizes, die Bewegungszeit für die resultierende Reaktion und die Zeit, die das bediente Gerät braucht, um die Reaktionshandlung zu verarbeiten. Die mentale Verarbeitungszeit kann jedoch ohne eine körperliche Reaktion nicht direkt gemessen werden.

In dieser Studie war die Reaktionszeit auf einen Stimulus die abhängige Variable in den Experimenten. Die Reaktionszeit wurde als die Zeit gemessen, die vom Auftauchen des Reizes in einer sichtbaren Entfernung von etwa 800 m bis zum Erkennen des Reizes vergeht. Im Experiment wurde ein physikalisches Gerät zur Bestätigung der Objekterkennung eingesetzt, in Form des im Triebfahrzeug vorhandenen Signalhorns, welches zu betätigen war. Dies entspricht durchaus auch einem in der Realität üblichen Verhalten nach dem Erkennen eines Objekts im Gleis.

Die Teilnehmenden wurden instruiert, im Experiment auf jedes erkannte Objekt so schnell wie möglich zu reagieren. Es musste explizit nicht unterschieden werden, ob das Objekt potentiell gefährlich war. Dies ermöglichte es den Teilnehmenden, sich auf ihre sensorischen Fähigkeiten zu konzentrieren. Es musste keine weitere kognitive Verarbeitung des Reizes verbunden mit einer Entscheidung über die Reaktion erfolgen.

Auf der Grundlage der gegebenen Reaktionszeiten und Fahrgeschwindigkeiten können der Abstand des Zuges zum nächsten Würfel zum Zeitpunkt der Reaktion und die Zeitspanne zwischen dem Reagieren auf den erscheinenden Stimulus und dem Passieren des Stimulus durch den Zug bei angenommener gleichbleibender Geschwindigkeit berechnet werden. Da diese Werte jedoch mit der unabhängigen Variablen der Fahrgeschwindigkeit konfundiert sind, können sie lediglich zur ergänzenden Analyse durch deskriptive Statistik herangezogen werden.

## 7 Experimentalstudien

### 7.1 Versuchsumgebung

#### 7.1.1 Fahrsimulator BBI



Abbildung 6: Die Simulationsumgebung der BBI

Die BBI-Studie wurde in einem Fahrsimulator der TU Berlin durchgeführt (Abbildung 6). Der Führerstand des Fahrsimulator bildet hinsichtlich der Bedienelemente und der Anzeigen einen Führerstand der Baureihen 193 oder 248 (Vectron Dual Mode) mit Anpassungen aufgrund der vorhandenen Hardware und sonstigen Mittel nach. Der Führerstand entspricht den Anforderungen des European Driver's Desk (EUDDplus). Alle relevanten technischen sowie betrieblich notwendigen Anzeigen wurden über mehrere Touchdisplays dargestellt. Der Fahrsimulator wurde mit der Software Zusi 3 Professional betrieben. Die Software bietet verschiedene Streckenmodule und realistische Betriebsregeln. Die Darstellung erfolgt auf einem 32 Zoll UHD-Bildschirm. Die simulierte Umgebung umfasst zentrale Elemente einer realen Bahnstrecke. Betrieblich relevante akustische Signale wie Sifa- und PZB-Warnungen werden realistisch umgesetzt.

#### 7.1.2 Fahrsimulator DLR



Abbildung 7: Die RailSET Simulationsumgebung

Die DLR-Studie wurde im Triebfahrzeugsimulator RailSET® (Railway Simulation Environment for Train drivers and operators) des Instituts für Verkehrssystemtechnik des DLR durchgeführt (Johne und Busse,

2016). Der Simulator wurde mit Hilfe eines Original-Steuerpults eines Triebfahrzeugs bedient (Abbildung 7). Die Simulationsumgebung basiert auf der Software VIRES (VIRES Simulationstechnologie GmbH, Bad Aibling, Deutschland). Ein Videoprojektor zeigt die simulierte Sicht nach vorne, während die Sicht aus den seitlichen Fenstern auf Bildschirmen angezeigt wird. Ein Audiosystem in der Kabine liefert Umgebungsgeräusche, die dem Innenraum eines echten Triebfahrzeugs nachempfunden sind.

### 7.1.3 Onlinestudie

Eine Onlinestudie mit Laien ohne Tf-Ausbildung wurde durchgeführt, um die Stichprobengröße zu erhöhen und die Ergebnisse der Simulatorstudien zu ergänzen. In der Onlinestudie sahen die Teilnehmenden Videos, die im Simulator des DLR aus der Perspektive der Tf während der Annäherung an die Würfel als Screenvideo aufgezeichnet wurden. Die Teilnehmenden sollten mit einem Druck auf die Entertaste reagieren, sobald sie den Würfel wahrnahmen.

Die Onlinestudie wurde unter Verwendung des jsPsych-Pakets für die Präsentation der Videos und Limesurvey für die begleitenden Fragebögen durchgeführt (Leeuw, 2015). Für die Verwendung der Videos in einer Onlinestudie wurden diese im mp4-Format verwendet, das eine optimale Dateikomprimierung bei gleichbleibender Wiederholungsrate erlaubt. Zudem wurde die Auflösung auf 1280 × 720 Pixel reduziert. Bei der Auswertung der Onlinestudie sollte beachtet werden, dass dieser Auflösungsunterschied theoretisch zu einer schlechteren Sichtbarkeit führen kann; es gab darauf aber keine Kommentare oder Hinweise von den Probanden. Um sicherzustellen, dass die Videos reibungslos abgespielt wurden und die Reaktionszeitmessung nicht durch Ladeverzögerungen beeinträchtigt wurde, wurden vor Beginn des Experiments alle Videos vollständig geladen. Bei einigen Geräten war das Vorladen jedoch trotz einer schnellen und zuverlässigen Internetverbindung nicht möglich. Für diese Fälle wurde eine Umleitung eingerichtet, unter der das Experiment dann ohne Vorladen durchgeführt werden konnte. Ob Teilnehmende das Experiment im Standardmodus (mit Preload) oder im Fehlerzustand (ohne Preload) beendeten, wurde protokolliert, um die Daten auf systematische Unterschiede zu überprüfen.

## 7.2 Szenarien

Die Aufgabe der teilnehmenden Tf bestand darin, einen Zug zu fahren und dabei mit einem Tastendruck auf unregelmäßig erscheinende Stimuli in Form von Würfeln zu reagieren. Die Teilnehmenden befuhren drei verschiedene Strecken in drei Versuchsblöcken:

- Fahren mit dem Zugsicherungssystem PZB (PZB-Strecke),
- Fahren mit dem Zugsicherungssystem ETCS Level 2 (ETCS-Strecke) und
- Fahren auf Sicht (FaS).

Die in den BBI- und DLR-Studien verwendeten Strecken waren unterschiedlich, es wurde jedoch darauf geachtet, die Objekte an ähnlichen Stellen zu platzieren, z. B. auf geraden Strecken mit geringer oder keiner Neigung.

Für den Versuchsblock Fahren auf Sicht (FaS) wurden die Teilnehmenden instruiert, einen Teil der PZB-Strecke auf Sicht zu befahren. Fahren auf Sicht bedeutet, dass die Tf nicht von einem Freisein der Strecke angezeigt durch grüne Signale ausgehen können. Stattdessen haben sie die Aufgabe, mögliche Gefahren durch eine genaue Streckenbeobachtung zu identifizieren und wenn nötig zum Halten zu kommen. Beim Fahren auf Sicht können die Tf ihre Fahrgeschwindigkeit in Abhängigkeit von den Sicht- und

Gleisverhältnissen selbst wählen, diese darf aber 40 km/h nicht überschreiten. Dieses Szenario wurde als erste Fahrt nach einem Sturm beschrieben.

### 7.2.1 Szenarien BBI

Entlang der ca. 31,5 km langen PZB-Strecke wurden acht Stimuli platziert. Die ETCS-Strecke (50 km) umfasste größtenteils eine gerade Strecke mit der maximal zugelassenen Geschwindigkeit von 160 km/h. Entlang der ETCS-Strecke wurden fünf Stimuli platziert. Die Unterschiede in der Anzahl der Stimuli zwischen den Szenarien sind durch die Randbedingungen des Simulatoreaufbaus entstanden. Die Fahrzeiten in den PZB- und ETCS-Szenarien lagen bei etwa 25-30 Minuten. Die FaS-Strecke war kürzer (6 km, ca. 10-15 Minuten) und umfasste vier Stimuli mit unterschiedlicher Größe und Kontrast. Beim Fahren auf Sicht (FaS) fuhren die Tf den Zug ohne Unterstützung der Zugsicherungssysteme mit einer niedrigeren Geschwindigkeit und mussten diese je nach Strecke und Sichtverhältnissen anpassen.

### 7.2.2 Szenarien DLR

Die ETCS-Strecke ist, ähnlich einem typischen Abschnitt einer Hochgeschwindigkeitsstrecke in Deutschland, eine weitgehend gerade Strecke mit wenigen, weiten Kurven. Entlang der ETCS-Strecke wurden acht Stimuli platziert. Die PZB-Strecke ist deutlich kurviger als die ETCS-Strecke. Die zulässige Höchstgeschwindigkeit variiert häufig. Genau wie bei der ETCS-Strecke wurden acht Stimuli entlang der PZB-Strecke platziert. Es dauerte jeweils ca. 35 min, die ETCS-Strecke (100 km) bzw. die PZB-Strecke (31 km) zu befahren. Beim Fahren auf Sicht wurden zwei große Stimuli platziert. Das Platzieren einer höheren Anzahl an Stimuli war aufgrund der niedrigeren Geschwindigkeit und somit einer längeren Fahrzeit beim Fahren auf Sicht nicht realisierbar. Deshalb wurde hier in Bezug auf die physikalischen Eigenschaften der Stimuli nur der Kontrast variiert, nicht aber die Stimulusgröße. Das Fahren auf Sicht dauerte ca. 10 min (5 km).

### 7.2.3 Szenarien der Online-Studie

Ziel der Online-Studie war es, den Datensatz aus der Simulatorstudie durch eine größere Stichprobe zu ergänzen. Der angestrebte Personenkreis für die Online-Studie waren dabei Laien, d. h. Personen ohne eine Tf-Ausbildung. Zu diesem Zweck wurden in der DLR-Simulationsumgebung Videos von der Annäherung an jeden Stimulus auf der Strecke in der Simulationsumgebung aufgezeichnet. Diese Videos wurden dann in eine Onlineumfrage eingebettet und den Teilnehmenden der Online-Studie in zufälliger Reihenfolge dargeboten. Äquivalent zur Simulatorstudie hatten die Teilnehmenden die Aufgabe, mit einem Tastendruck auf die "Enter"-Taste auf der Tastatur zu reagieren, sobald sie den Stimulus wahrgenommen hatten. Im Rahmen der Onlinestudie wurde FaS nicht umgesetzt.

## 7.3 Studiendesign

Die Studie war als Within-Subjects-Design angelegt, bei dem jede Versuchsbedingung durch einen Stimulus während der Fahraufgabe repräsentiert wurde. Jeder Teilnehmende bearbeitete jede Versuchsbedingung. Die untersuchten unabhängigen Variablen lassen sich zwei Kategorien zuordnen:

- 1) physikalische Eigenschaften der Stimuli
- 2) betriebliche Parameter.

Für eine Erläuterung der unabhängigen Variablen siehe Kapitel 6.3. In Tabelle 8 sind die Versuchsbedingungen als Übersicht dargestellt.

Die unabhängigen Variablen der Onlinestudie waren bis auf eine Ausnahme die gleichen wie in der Simulatorstudie. Es wurden die Größe und der Kontrast des Stimulus variiert, genauso wie die Zuggeschwindigkeit. Anders als in der Simulatorstudie gab es keine Bedingung mit Fahrt auf Sicht, es wurden aber Fahrten mit PZB sowie mit ETCS angeboten. Es ist jedoch anzumerken, dass sich die erwartete Wirkweise des Faktors des Zugsicherungssystems in der Online-Studie grundlegend von der erwarteten Wirkweise des Faktors der Simulatorstudie unterscheidet. Während PZB und ETCS in der Simulatorstudie die Aufgaben der Tf und den Umfang der verfügbaren visuellen Aufmerksamkeit verändern, mussten die Teilnehmenden der Online-Studie nur Videos von der Fahrt ansehen. Es entstanden keine zusätzlichen Aufgaben durch PZB oder ETCS. Daher ist die Relevanz des Zugsicherungssystems in der Online-Studie eher im unterschiedlichen Streckendesign der beiden Streckenabschnitte zu sehen. Wie bei der Simulatorstudie war die abhängige Variable der Online-Studie die Reaktionszeit, d. h. die Zeit zwischen dem Erscheinen des Stimulus 800 m vor dem Zug und der Reaktion durch Tastendruck.

TABELLE 8: EXPERIMENTALBEDINGUNGEN DER DREI STUDIEN.

	40 km/h		100 km/h		160 km/h	
<b>FaS (BBI)</b>	kontrastreich	groß				
		klein				
	kontrastarm	groß				
		klein				
<b>FaS (DLR)</b>	kontrastreich	groß				
	kontrastarm	groß				
<b>PZB (BBI + DLR + Online)</b>	kontrastreich	groß	kontrastreich	groß		
		klein		klein		
	kontrastarm	groß	kontrastarm	groß		
		klein		klein		
<b>ETCS (BBI)</b>			kontrastreich	klein	kontrastreich	klein
			kontrastarm	groß klein		
<b>ETCS (DLR + Online)</b>			kontrastreich	groß	kontrastreich	groß
				klein		klein
			kontrastarm	groß	kontrastarm	groß
				klein		klein

## 7.4 Forschungshypothesen

In diesem Abschnitt werden die Forschungshypothesen vorgestellt, die die Analyse und Untersuchung der Beziehung zwischen verschiedenen Faktoren und der Reaktionszeit (RT) leiten werden.

## 7.4.1 Simulatorstudien

Alle Hypothesen wurden mit einseitigen t-Tests bei einem Konfidenzniveau von  $\alpha = 0,05$  getestet. Es wurde erwartet, dass ein geringerer Kontrast und eine geringere Größe die Auffälligkeit der Stimuli verringern. Daher sollten kleinere Stimuli später erkannt werden als große Stimuli und auf kontrastarme Stimuli sollte langsamer reagiert werden als auf kontrastreiche Stimuli. In Bezug auf die Parameter  $\beta_1, \dots, \beta_{4.2}$  des linearen Regressionsmodells (Abschnitt 7.9) können diese Hypothesen wie folgt definiert werden:

H1: Die Reaktionszeit ist bei kleinen Stimuli länger:  $\beta_1 > 0$ .

H2: Die Reaktionszeit ist bei kontrastarmen Reizen länger:  $\beta_2 > 0$ .

Es wurde festgestellt, dass mit steigender Geschwindigkeit die Vertikalität der Augenbewegungen zunimmt und die Horizontalität der Augenbewegungen abnimmt (Suzuki, Yamauchi & Matsuura, 2019). Das heißt, dass sich bei schnelleren Geschwindigkeiten die Aufmerksamkeit der Tf stärker geradeaus entlang des Gleises orientiert. Da die Stimuli auf/direkt neben den Gleisen erschienen, sollte die stärkere Aufmerksamkeitsfokussierung in den Gleisbereich bei höheren Geschwindigkeiten dazu führen, dass auf Stimuli, die bei höheren Geschwindigkeiten erscheinen, schneller reagiert wird als auf solche bei niedrigeren Geschwindigkeiten.

H3a: Die Reaktionszeit ist bei langsamer Fahrgeschwindigkeit (40 km/h) länger als bei den Bedingungen mit hoher Fahrgeschwindigkeit:  $\beta_{3.1} > 0$ .

H3b: Die Reaktionszeit ist bei einer Fahrgeschwindigkeit von 100 km/h länger als bei 160 km/h:  $\beta_{3.2} > 0$ .

Beim Fahren auf Sicht verwenden die Tf den größten Teil ihrer Aufmerksamkeit auf die Beobachtung der Gleise. Wenn ein Zugsicherungssystem verwendet wird, müssen zusätzlich die Anzeigen im Führerstand beobachtet werden, d. h. die Aufmerksamkeit wird geteilt. Aus diesem Grund wurde erwartet, dass die Verwendung eines Zugsicherungssystems (PZB oder ETCS) zu längeren Reaktionszeiten führt als beim Fahren auf Sicht. Wie in Kapitel 6.3.2 erwähnt, führt die Aufteilung der visuellen Aufmerksamkeit im ETCS mit Führerstands-signalisierung zu einer erheblichen Verringerung der Zeit für die Beobachtung der Strecke (Marinakos et al 2005; Van der Weide et al. (2017). Dies könnte sich negativ auf die Objekterkennung auswirken. Zudem wurde erwartet, dass die Tf bei der Fahrt mit ETCS mit Führerstands-signalisierung langsamer auf das Auftreten der Stimuli reagieren als bei der Fahrt mit PZB, da die Aufmerksamkeitsanforderungen auf dem Fahrerpult bei Fahrt mit ETCS höher sind als bei Fahrt mit PZB.

H4a: Die Reaktionszeit ist länger, wenn ein Zugsicherungssystem verwendet wird als beim Fahren auf Sicht:  $\beta_{4.1} > 0$ .

H4b: Die Reaktionszeit ist länger, wenn mit ETCS gefahren wird als mit PZB:  $\beta_{4.2} > 0$ .

## 7.4.2 Onlinestudie

In Bezug auf die Regressionskoeffizienten des linearen Modells wurden, in Anlehnung an die Simulatorstudie, die folgenden Hypothesen definiert. H4 unterscheidet sich, da das Zugsicherungssystem hier keine zusätzliche Aufgabe bedeutete, sondern angenommen wurde, dass die abwechslungsreichere PZB-Route zu mehr visueller Ablenkung führt:

H1: Die Reaktionszeit ist bei kleinen Stimuli länger:  $\beta_1 > 0$ .

H2: Die Reaktionszeit ist bei kontrastarmen Reizen länger:  $\beta_2 > 0$ .

H3a: Die Reaktionszeit ist bei langsamer Fahrgeschwindigkeit (40 km/h) höher als bei den Bedingungen mit länger Fahrgeschwindigkeit:  $\beta_{3.1} > 0$ .

H3b: Die Reaktionszeit ist bei einer Fahrgeschwindigkeit von 100 km/h länger als bei 160 km/h:  $\beta_{3.2} > 0$ .

H4: Die Reaktionszeit ist kürzer, wenn mit ETCS gefahren wird als mit PZB:  $\beta_4 > 0$ .

## 7.5 Ablauf der Datenerhebung

Bei der Ankunft an den Simulatorstandorten des DLR und des BBI wurden die Teilnehmenden zunächst gebeten, die Erklärungen zum Datenschutz sowie die Probandeninformation zu lesen sowie die Einwilligungserklärung zu unterschreiben. Nachfolgend wurde allen Teilnehmenden der Ablauf der Datenerhebung erläutert. Die Teilnehmenden erhielten eine Einführung in die Fahraufgabe, wurden instruiert, bei Stimuluserscheinen das Signalhorn zu betätigen und erhielten eine kurze Einführung in das ETCS, da die meisten Teilnehmenden keine vorherige Arbeitserfahrung mit dem System hatten. Anschließend hatten die Teilnehmenden die Möglichkeit, Fragen zu stellen.

Bevor die Durchführung der Versuchsblöcke begann, füllten die Teilnehmenden auf einem Tablet den Fragebogen zu demografischen Merkmalen und die Karolinska-Sleepiness-Scale (Shahid et al. (2012)) aus. Anschließend wurde ein dreiminütiger Psychomotorischer Wachsamkeitstest (PVT) (Dinges und Powell, 1985) durchgeführt. Dieser Test besteht aus einem einfachen Reaktionstest auf ein Licht, welches in unregelmäßigen Abständen aufleuchtet. Ziel des Tests war es zu messen, ob über die Zeit, unabhängig von den Versuchsbedingungen, Müdigkeit und damit längere Reaktionszeiten aufkamen. Im Folgenden absolvierten die Teilnehmenden die drei Versuchsblöcke (ETCS-Strecke, PZB-Strecke und Fahren auf Sicht), wobei nach jedem Block die Karolinska-Sleepiness-Scale ausgefüllt wurde. Außerdem wurde der PVT vor jedem Lauf wiederholt. Die PZB- und ETCS-Strecken wurden immer als erstes absolviert, die Reihenfolge dieser Strecken war dabei zwischen den Teilnehmenden ausgeglichen. Die Fahrt auf Sicht wurde immer als letztes absolviert. Den Teilnehmenden wurde erklärt, dass sie das Signalhorn ziehen oder drücken sollten, sobald sie einen orangen oder braunen Würfel sahen. Es wurde erklärt, dass es keine Kollision geben würde und keine Notwendigkeit bestand, den Fahrbetrieb aufgrund der gezeigten Objekte zu ändern. Für die Durchführung der 3 Fahrten und das Ausfüllen der Fragebögen wurden ca. 3 Stunden pro Teilnehmendem benötigt.

Um an dem Onlineexperiment teilzunehmen, erhielten die Teilnehmenden zunächst Informationen über die Studie in Textform. Dann wurde die Zustimmung zur Datenerhebung abgefragt. Es folgten ein kurzer demographischer Fragebogen sowie die Erhebung der subjektiven Müdigkeit. Anschließend wurden die Teilnehmenden zu einem Server weitergeleitet, auf dem das eigentliche Experiment stattfand, d. h. die Videos präsentiert wurden. Im Rahmen des Experiments durchliefen die Teilnehmenden zunächst 3 Übungsrunden, dann folgten die 16 Bedingungen des Experiments. Jedes Video wurde von den Teilnehmenden manuell durch einen Druck auf die Leertaste gestartet. Die Teilnehmenden hatten die Aufgabe, die "Enter"-Taste auf ihrer Tastatur zu drücken, sobald sie einen Stimulus auf den Gleisen sahen. Die Stimuli erschienen zwischen 12 und 70 Sekunden nach Beginn des Videos. Der Zug passierte den Stimulus zwischen 46 und 133 Sekunden nach dem Start des Videos, je nach Annäherungsgeschwindigkeit und Erscheinungszeit.

## 7.6 Stichprobe

Es wurden ausgebildete Tf für die Teilnahme an der Simulatorstudie rekrutiert. Die Rekrutierung erfolgte über den Teilnehmendenpool des DLR und der TU Berlin sowie über Werbung für die Studie in sozialen Medien. Die Wahl aktiver Tf als Teilnehmende stellte sicher, dass die Ergebnisse auf realen beruflichen Erfahrungen und Fähigkeiten basieren.

An der BBI-Studie nahmen 18 aktive Tf mit einem Durchschnittsalter von 33,4 Jahren (Spanne = 22-57 Jahre) teil. Die Berufserfahrung der Teilnehmenden lag zwischen 1 und 28 Jahren (Mittelwert = 7,3). Die Teilnehmenden bewerteten den Vertrautheitsgrad mit verschiedenen Arten von Zugsicherungssystemen auf einer Skala von 1 (überhaupt nicht vertraut) bis 10 (sehr vertraut). Die durchschnittliche Vertrautheit mit dem PZB-System wurde mit einem Durchschnitt von 8,3 bewertet, während die Vertrautheit mit dem ETCS einen Durchschnitt von 2,2 erreichte.

An der DLR-Studie nahmen insgesamt 25 Berufstriebfahrzeugführer mit im Schnitt 9,92 Jahren Berufserfahrung (Spanne = 1-39 Jahre) teil. Die Teilnehmenden waren alle männlich und im Schnitt 33,7 Jahre alt (Spanne = 22-57 Jahre). Keiner der Teilnehmenden hatte zuvor an einer Studie im RailSET teilgenommen. Auf einer Skala von 1 (überhaupt nicht vertraut) bis 10 (sehr vertraut) bewerteten die Teilnehmenden ihre Vertrautheit mit dem Zugsicherungssystem PZB im Schnitt mit 9,8, während die Vertrautheit mit dem Zugsicherungssystem ETCS im Durchschnitt mit 2,0 bewertet wurde (Tabelle 9).

TABELLE 9: STICHPROBEN IN DEN EXPERIMENTELLEN STUDIEN

	<b>BBI</b>	<b>DLR</b>	<b>Online</b>
Probanden	18 Tf	25 Tf	70 Nicht-Tf
Alter (Mittelwert)	22-57 (33,4)	22-57 (33,7)	19-65 (31,9)
Berufserfahrung in Jahren (Mittelwert)	1-28 (7,3)	1-39 (9,9)	-
PZB-Vertrautheit	8,3/10	9,8/10	-
ETCS-Vertrautheit	2,2/10	2,0/10	-

An der Onlinestudie konnte jeder teilnehmen, der keine Ausbildung als Tf hatte sowie über eine stabile Internetverbindung und einen Computer mit Tastatur verfügte. Die Teilnehmenden waren 70 Nicht-Tf, die hauptsächlich aus den Probandenpools des DLR und des BBI rekrutiert wurden. Die Teilnehmenden waren im Schnitt 31,95 Jahre alt (Spanne = 19-65 Jahre). Aufgrund von technischen Schwierigkeiten mit dem Onlinefragebogen fehlten bei fünf der Teilnehmenden demografische Daten. Von den übrigen 65 waren 34 weiblich und 31 männlich. Die Reaktionszeitdaten der Teilnehmenden mit fehlenden demografischen Daten wurden in die Analyse einbezogen, da keine Unregelmäßigkeiten in den Experimentaldaten auftraten.

## 7.7 Datenvorverarbeitung

Zunächst wurden aus den Protokollen des Simulators die Zeitpunkte des Auftretens der einzelnen Stimuli ausgelesen. Zudem wurden die Zeitpunkte ausgelesen, zu denen die Taste für das Signalhorn betätigt wurde. Anhand der Differenz beider Werte wurde für jede Stimuluspräsentation die Reaktionszeit berechnet. Da das Signalhorn auch im normalen Zugbetrieb verwendet wird, z. B. vor einer scharfen Kurve, drückten einige Teilnehmende die Taste auch dann, wenn kein Stimulus in Sicht war. Daher wurden alle Tastendrucke außerhalb des Erscheinungsfensters der Stimuli von der Analyse ausgeschlossen. In Fällen, in denen die Taste für das Signalhorn bei Entdecken des Stimulus mehrmals betätigt wurde, wurde das erste Drücken gezählt. Um sicherzustellen, dass die gefahrene Geschwindigkeit zum Zeitpunkt des Erscheinens des Würfels auch der Geschwindigkeit entsprach, die vorgegeben war (also 40 km/h, 100 km/h oder 160 km/h), wurde eine Analyse der Abweichung zwischen vorgegebener und tatsächlich gefahrener Geschwindigkeit durchgeführt. Es wurde der Interquartilabstand der Verteilung der tatsächlich gefahrenen Geschwindigkeit erstellt. Es wurden alle Datenpunkte aus der Analyse ausgeschlossen, bei denen die tatsächlich gefahrene Geschwindigkeit beim Erscheinen des Stimulus stärker als der dreifache Interquartilabstand von der vorgegebenen Geschwindigkeit abwich (sechs Fälle beim DLR, fünf Fälle beim BBI). Obwohl das 1,5-fache des IQR in der Literatur am häufigsten verwendet wird (Tukey, 1977), kann das 3-fache zur Ermittlung extremer Ausreißer verwendet werden, um eine weniger konservative Cut-off-Methode anzuwenden. Bei Datensätzen mit geringem Stichprobenumfang stellt die Methode mit 3IQR sicher, dass nur extrem abweichende Werte als Ausreißer betrachtet werden. Von 306 Rohdatenpunkten der BBI-Studie wurden 19 Datenpunkte aufgrund falscher Reaktionszeitstempel oder technischer Probleme während der Datenerfassung eliminiert. Insgesamt lagen nach Datenbereinigung 426 Beobachtungen von 25 Teilnehmenden des DLR und 282 Beobachtungen von 18 Teilnehmenden der BBI-Studie vor. Die Abbildung 8 zeigt die Anzahl der Beobachtungen unter verschiedenen Versuchsbedingungen für die Simulatorstudien.

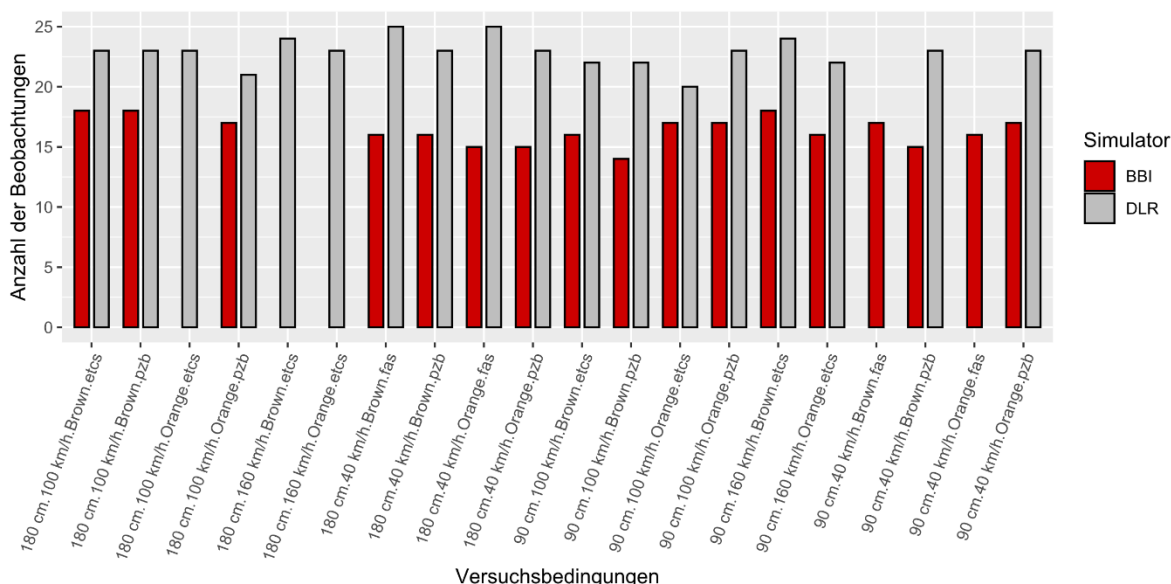


Abbildung 8: Anzahl der Datenpunkte (y-Achse) für jede Bedingung als Objektgröße, Geschwindigkeit, Objektfarbe, Zugsicherungssystem (x-Achse)

Auch bei der Online-Studie wurde ein Datenverarbeitungsverfahren angewandt. Nach der Datenerhebung wurden die Daten für die Fragebögen und die Reaktionszeitdaten zunächst auf der Grundlage einer laufenden Probanden-ID zusammengefasst. In den Fragebögen gaben einige Teilnehmende technische Schwierigkeiten wie z. B. sehr lange Ladezeiten oder einen leeren Bildschirm anstelle eines Videos an. Wenn fehlerhafte Bedingungen anhand der Beschreibung der Teilnehmenden identifiziert wurden, wurden nur diese Bedingungen in den Daten des betreffenden Teilnehmenden ausgeschlossen (neun Datenpunkte). Wenn es unklar war, welche Bedingungen betroffen waren oder das Problem als durchgängig beschrieben wurde, wurden die Daten des betreffenden Teilnehmenden komplett ausgeschlossen (zwei Datensätze). Abschließend wurden frühe Tastendrucke vor Erscheinen des Stimulus sowie verpasste Stimuli, bei denen der Tastendruck erst nach dem Passieren des Stimulus erfolgte, identifiziert und aus dem Datensatz ausgeschlossen (154 Datenpunkte). Insgesamt ergab die Datenbereinigung einen Satz von 925 Beobachtungen von 68 Teilnehmenden.

## 7.8 Statistische Auswertung

In Bezug auf die Reaktionszeitdaten wurde im ersten Schritt eine deskriptive Analyse der Reaktionszeitdaten als notwendige Vorstufe zu einer inferenzstatistischen Analyse durchgeführt. Reaktionszeiten haben eine Verteilung mit einer harten unteren Grenze von 0 Sekunden und weisen meist eine große Anzahl sehr schneller Reaktionen sowie einen langen Verteilungsenden mit hohen Reaktionszeiten auf. In beiden Studien wurden positiv verzerrte Reaktionszeiten festgestellt. Das Histogramm der Reaktionszeiten ist in Abbildung 9 (links) zu sehen. Das Cullen-Frey-Diagramm zeigt die Schiefe und die Kurtosis der Daten, zwei Werte, die die Form der Verteilung beschreiben. Der Skewness-Wert deutet auf einen längeren rechten Ausläufer und eine positive Schiefe im Vergleich zu einer symmetrischen Verteilung hin (Abbildung 9, rechts). Der Kurtosis-Wert von 10,6 könnte auf eine Verteilung mit mehr extremen Werten im Vergleich zu einer Normalverteilung hindeuten. Daher sollte vor der Datenanalyse eine Transformationsmethode angewandt werden.

Die Transformation der Reaktionszeiten ist eine etablierte Methode, um statistische Tests mit Modellen zu ermöglichen, die eine Normalverteilung der Datenpunkte oder der Fehler (Residuen) annehmen (Czamolewski, 1996). Durch die Anwendung der logarithmischen Transformation haben Ausreißer einen vergleichsweise geringeren Einfluss auf die Analyse (Baayen und Milin, 2010). Es gibt mehrere Ansätze, um zu bestimmen, welches theoretische Modell die Daten am besten beschreibt. Eines der Kriterien ist, wie praktisch es ist, die Daten zurück zu transformieren, um die Auswirkungen auf die ursprüngliche Reaktionszeitskala darzustellen (Baayen und Milin, 2010). Aus diesen Gründen wurden die Reaktionszeiten für das Regressionsmodell logarithmisch transformiert.

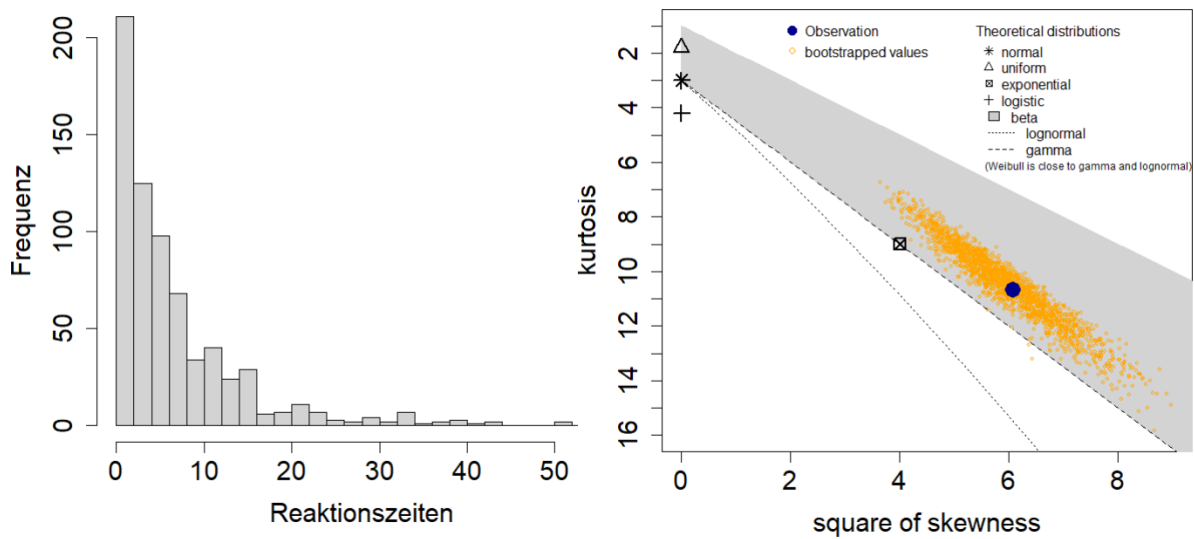


Abbildung 9: Histogramm der Reaktionszeitdaten (links) und das Cullen-Frey-Diagramm (rechts)

Drittens wurde die Analyse des Einflusses der unabhängigen Variablen anhand eines linearen Regressionsmodells vorbereitet. Für die Berechnung eines linearen Regressionsmodells gibt es verschiedene Annahmen, die von den Daten erfüllt werden müssen (Cohen und Cohen, 2013). Die erste dieser Annahmen ist die Normalverteilung der Residuen. Es wurde mittels visueller Inspektion festgestellt, dass die Residuen annähernd normalverteilt sind, wenn die Reaktionszeiten log-transformiert wurden (vgl. Abbildung 10).

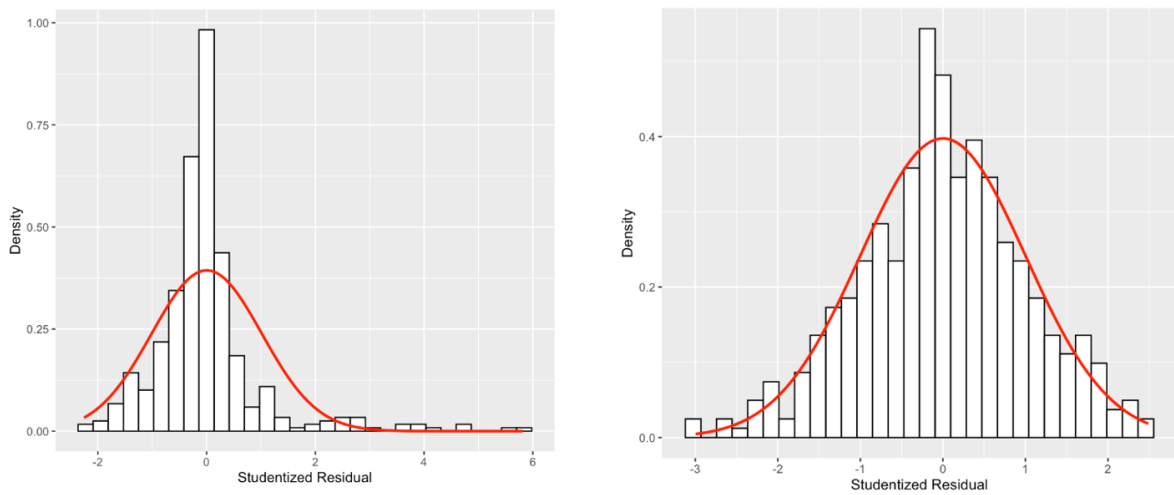


Abbildung 10: Residuen des rohen Reaktionszeitmodells (links) und des log-transformierten Reaktionszeitmodells (rechts) für die DLR-Studie verglichen.

Als zweite Annahme wurden die Anzahl und der Einfluss großer standardisierter Residuen geprüft. Das Modell wies eine akzeptable Anzahl großer standardisierter Residuen auf (22 Fälle in 426 Beobachtungen), die alle einen geringen Einfluss auf das endgültige Modell hatten. Daher wurden diese Fälle in das endgültige Modell aufgenommen. Die Annahme der Unabhängigkeit der Residuen wurde mit dem Durbin-Watson Test überprüft, der nur eine geringe positive Autokorrelation auf einem akzeptablen Niveau ergab (D-W-Statistik = 1,65). Die Teststatistik kann zwischen 0 und 4 liegen, wobei ein Wert von 2 bedeutet, dass die Residuen unkorreliert sind. Die Multikollinearität wurde durch Prüfung des Variance In-

flation Factor und der Toleranz bewertet, die sich in einem akzeptablen Bereich bewegten (DLR: mittlerer VIF = 1,55, min Toleranz > 0,3. BBI: mittlerer VIF = 1,95, min Toleranz > 0,34). Ein VIF höher als 4 und eine Toleranz weniger als 0,25 deuten auf Multikollinearität hin.

In der Onlinestudie wurde eine Kurzanalyse nach dem Vorbild der Simulatorstudie durchgeführt. Müdigkeit und demografische Daten wurden deskriptiv ausgewertet. Die Reaktionszeitdaten wurden zunächst ebenfalls deskriptiv ausgewertet. Darauf folgend wurde eine Regressionsanalyse durchgeführt. Aufgrund der nicht-normalen Verteilung wurden die Hypothesen auf geometrischen Mittelwerten unter Verwendung eines linearen Regressionsmodells mit log-transformierten Daten getestet. Die notwendigen Annahmen für die Berechnung einer Regressionsanalyse wurden geprüft. Anhand einer visuellen Inspektion wurde auf Normalverteilung der Residuen getestet. Wie in der Simulatorstudie zeigten die Residuen eine annähernd normale Verteilung, wenn die Reaktionszeiten logarithmisch transformiert wurden. Als zweite Annahme wurden die Anzahl und der Einfluss großer standardisierter Residuen geprüft. Bei einer Normalverteilung sollten 95 % der Daten innerhalb von 2 Standardabweichungen vom Mittelwert liegen (Ziegel und Ott, 1994). Das Modell wies eine akzeptable Anzahl großer standardisierter Residuen auf (42 Fälle in 925 Beobachtungen), die alle einen geringen Einfluss auf das endgültige Modell hatten. Daher wurden diese Fälle in das endgültige Modell aufgenommen. Die Multikollinearität wurde durch Prüfung des VIF und der Toleranz bewertet, sich in einem akzeptablen Bereich bewegten (mittlerer VIF = 1,30, minimale Toleranz > 0,49).

## 7.9 Definition des Regressionmodells

Im Folgenden werden die Schritte zur Definition des zu testenden linearen Regressionsmodells beschrieben.

### 7.9.1 Simulatorstudien

Ein erster Modellvergleich zwischen einem reinen Intercept-Modell (ein Modell, welches nur den Achsenabschnitt einbezieht) und einem zufälligen Intercept-Modell (ein Modell, welches eine Variation der Achsen über Versuchspersonen zulässt) mittels Maximum-Likelihood-Schätzung ergab, dass die Einbeziehung der Probanden-ID (d. h. die Identifizierung der Teilnehmenden) das Modell nicht signifikant verbesserte (L-Ratio = 1,35,  $p = 0,25$ ). Daher wurde für die Analyse ein Regressionsmodell ohne Berücksichtigung der Probanden-ID gewählt.

Das Modell umfasste alle unabhängigen Variablen als Faktoren, wobei eine lineare Beziehung der folgenden Form angenommen wurde:

$$\text{Log(Reaktionszeit)} = \beta_0 + \beta_1 \cdot \text{size} + \beta_2 \cdot \text{contrast} + \beta_{3,1} \cdot \text{speed}_1 + \beta_{3,2} \cdot \text{speed}_2 + \beta_{4,1} \cdot \text{zss}_1 + \beta_{4,2} \cdot \text{zss}_2 + \varepsilon \quad (\text{G1})$$

Die unabhängigen Variablen umfassen Größe, Kontrast, Geschwindigkeit und Zugsicherungssystem. In diesem Zusammenhang steht  $\beta_0$  für den Achsenabschnitt (Intercept) des Modells, während  $\beta_1$  bis  $\beta_{4,2}$  die unabhängigen Variablen repräsentieren und somit die Parameter darstellen, die die Steigung der Regressionsgeraden beeinflussen. Interaktionen wurden nicht in das Modell aufgenommen, da sich die Hypothesen auf die Signifikanz der Haupteffekte konzentrierten. Der Fehlerterm  $\varepsilon$  steht für unerklärte Variation.

Im Folgenden wird beschrieben, wie die unabhängigen Variablen im Rahmen des Modells codiert wurden. Die zweistufigen Faktoren Größe und Kontrast des Stimulus wurden mit einer binären Dummy-Variablen codiert, während die dreistufigen Faktoren Geschwindigkeit (40 km/h, 100 km/h, 160 km/h) und Zugsicherungssystem (Fahren auf Sicht, PZB, ETCS) als Kontraste definiert wurden (Tabelle 10). Kontraste sind für Variablen mit mehr als zwei Stufen in einer linearen Regression erforderlich und bestimmen, wie die Bedingungen miteinander verglichen werden. Da maximal zwei Vergleiche für einen dreistufigen Faktor durchgeführt werden können, wurden Vergleiche priorisiert. Bei den Dreistufenvariablen werden zwei Stufen zusammengefasst und mit der verbleibenden Stufe durch Helmert-Kodierung verglichen. Das bedeutet, dass zum Beispiel Fahren auf Sicht mit dem gemeinsamen Durchschnitt von PZB und ETCS verglichen wird (zss1). Eine zweite Variable vergleicht dann PZB und ETCS (zss2). Dabei muss die Summe der Kontraste immer null ergeben (siehe Werte für zss1 und zss2 Kontraste). In Bezug auf das Zugsicherungssystem wurde der Betrieb ohne Zugsicherungssystem (Fahren auf Sicht) mit dem Betrieb mit Zugsicherungssystem verglichen (zss1-Kontrast). Die Fahrten mit PZB und ETCS wurden mit dem Kontrast zss2 verglichen. In Bezug auf die Geschwindigkeit wurde die langsame Bedingung bei 40 km/h mit dem Mittelwert der beiden schnellen Bedingungen über den Kontrast speed1 verglichen, und die beiden schnelleren Bedingungen mit 100 km/h bzw. 160 km/h wurden über den Kontrast speed2 verglichen.

TABELLE 10: KONTRASTE UND DUMMY-KODIERUNG FÜR ALLE FAKTOREN DER SIMULATORSTUDIEN

Zugsicherungssystem	zss1	zss2	Geschw.	speed1	speed2	Größe	size	Kontrast	contrast
FaS	-2	0	40 km/h	2	0	groß	0	kontrastreich	0
PZB	1	-1	100 km/h	-1	1	klein	1	kontrastarm	1
ETCS	1	1	160 km/h	-1	-1				

## 7.9.2 Onlinestudie

Im Gegensatz zur Simulatorstudie zeigte der erste Modellvergleich anhand eines Likelihood ratio-tests zwischen einem reinen Intercept-Modell und einem Random-Intercept-Modell bei der Onlinestudie einen klaren Vorteil der Einbeziehung der Probanden-ID als Random-Effect ( $p < 0,01$ ). Daher folgte die Analyse für die Onlinestudie einem mehrstufigen Modellierungsansatz, bei dem die einzelnen Stimulusbeobachtungen als Variablen der Stufe 1 und die Probanden-ID als Variablen der Stufe 2 genutzt wurden.

In das Modell wurden die unabhängigen Variablen als Faktoren einbezogen, wobei eine lineare Beziehung angenommen wurde:

$$\log(\text{Reaktionszeit})_i = (\alpha + \mu_0) + \beta_1 \cdot \text{size} + \beta_2 \cdot \text{contrast} + \beta_{3.1} \cdot \text{speed}_1 + \beta_{3.2} \cdot \text{speed}_2 + \beta_{4.1} \cdot \text{zss}_1 + \beta_{4.2} \cdot \text{zss}_2 \quad (G 2)$$

In dieser Formel stellt  $\mu_0$  die Differenz des Achsenabschnitts für jeden Teilnehmenden dar. Die berücksichtigten Faktoren wurden größtenteils auf die gleiche Weise codiert wie in der Analyse der Simulatorstudie. Eine Ausnahme bildete das Zugsicherungssystem, da dieser Faktor in der Onlinestudie nur zwei Stufen hatte. Die verwendeten Kontraste und die Dummy-Kodierung für zweistufige Faktoren sind in Tabelle 11 dargestellt.

TABELLE 11: KONTRASTE UND KODIERUNG VON DUMMY-VARIABLEN FÜR ALLE FAKTOREN

Speed	speed1		speed2			
40 km/h	2		0			
100 km/h	-1		1			
160 km/h	-1		-1			
ZSS	zss	Size	size	Contrast	contrast	
PZB	1	klein	1	niedrig	1	
ETCS	0	groß	0	hoch	0	

## 8 Ergebnisse Simulatorstudie BBI

### 8.1 Deskriptive Analyse

Es wurde eine deskriptive Analyse durchgeführt, um die Daten zu charakterisieren. Dazu gehörten u. a. die Berechnung von Mittelwerten, Standardabweichungen und Korrelationskoeffizienten. Abbildung 11 stellt die Maße der zentralen Tendenz auf dem Streudiagramm der Reaktionszeit in Abhängigkeit von der Geschwindigkeit und der Objektgröße bei verschiedenen Zugsicherungssystemen dar. Die arithmetischen Mittelwerte sind über alle Bedingungen zu höheren Reaktionszeiten verzerrt.

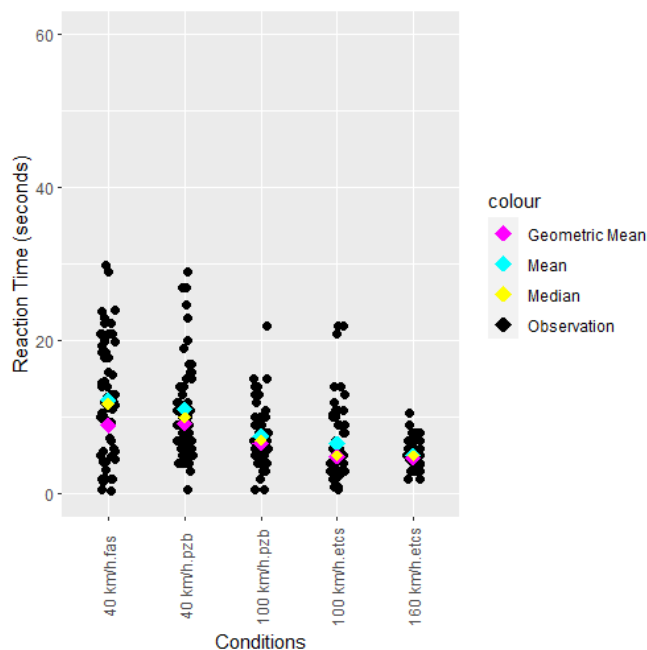


Abbildung 11: Maße der zentralen Tendenz für einige Stimulus-Eigenschaftsbedingungen. Geometrische Mittelwerte sind in magenta, arithmetische Mittel in blau und Medianwerte in gelb.

Die Tabelle 12 enthält die geometrischen Mittelwerte der Reaktionszeiten für jeden Stimulus. Generell führen höhere Geschwindigkeiten zu kürzeren Reaktionszeiten, was auf eine schnellere Reaktion auf Reize hinweist. Es ist auch ein Faktor, dass bei erhöhter Geschwindigkeit Objekte visuell schneller an Größe zunehmen und deshalb auch schneller erkannt werden können. Hoher Kontrast führt in der Regel zu kürzeren Reaktionszeiten im Vergleich zu niedrigem Kontrast. Auch die Objektgröße spielt eine Rolle, wobei kleinere Objekte in der Regel mit längeren Reaktionszeiten verbunden sind. Bei den meisten FaS-Bedingungen sind die Reaktionszeiten kürzer als bei gleichen Bedingungen mit PZB. Im Gegensatz zu den Erwartungen führte das ETCS-Szenario zu kürzeren Reaktionszeiten im Vergleich zur PZB. Der Korrelationskoeffizient zwischen Mittelwert und Standardabweichung der Reaktionszeiten unter verschiedenen Bedingungen beträgt 0,88. Diese positive Korrelation zwischen Reaktionszeiten und Standardabweichungen deutet darauf hin, dass Bedingungen, die längere Reaktionszeiten hervorrufen, tendenziell auch eine größere Variabilität in den Reaktionszeitmessungen aufweisen. Die spezifischen Auswirkungen jeder Variable können mithilfe Regressionsmodell weiter analysiert werden, um ihre Signifikanz zu bestimmen und die Stärke ihres Einflusses auf die Reaktionszeiten zu beurteilen.

TABELLE 12: GEOMETRISCHE MITTELWERTE UND GEOMETRISCHE STANDARDABWEICHUNGEN (IN KLAMMERN GESETZT) DER REAKTIONSZEIT DER VERSCHIEDENEN VARIABLENKOMBINATIONEN IN DER BBI-SIMULATORSTUDIE.

	40 km/h			100 km/h			160 km/h		
<b>FaS</b>	kontrast-reich	groß:	6,35 (2,59)						
		klein:	9,92 (1,89)						
	kontrast-arm	groß:	7,25 (4,21)						
		klein:	13,44 (1,66)						
<b>PZB</b>	kontrast-reich	groß	7,02 (1,67)	kontrast-reich	groß	5,99 (1,38)			
		klein	10,80 (1,74)		klein	5,79 (2,87)			
	kontrast-arm	groß	8,51 (1,60)	kontrast-arm	groß	6,80 (1,53)			
		klein	11,05 (2,77)		klein	8,07 (1,32)			
<b>ETCS</b>				kontrast-reich	klein	5,17 (2,07)	kontrast-reich	klein	4,60 (1,30)
				kontrast-arm	groß	3,41 (1,34)	kontrast-arm	klein	4,78 (1,70)
					klein	6,74 (3,24)			

## 8.2 Ergebnisse der Regressionsanalyse

Die Auswirkungen der unabhängigen Variablen auf die Reaktionszeit wurden mithilfe einer Regressionsanalyse analysiert. Die Ergebnisse, einschließlich der Koeffizienten des Modells und die Teststatistiken, sind in Tabelle 13 zu finden. Die Koeffizienten in der Tabelle entsprechen den Parametern des linearen Modells – z. B. entspricht  $\beta_1$  der Auswirkung des Stimulus Größe auf das Modell. Da die Daten in eine logarithmische Skala umgewandelt wurden, wurden auch exponentielle Koeffizienten berechnet. Die  $p$ Werte in der Tabelle geben das Signifikanzniveau für den einseitigen Test der Koeffizienten gegen Null an. Ein einseitiger  $p$ -Wert unterhalb des festgelegten Signifikanzniveaus (0,05) deutet darauf hin, dass der entsprechende Koeffizient statistisch signifikant und von Null verschieden ist.

TABELLE 13: REGRESSIONSKOEFFIZIENTEN UND T-WERTE FÜR LINEARE REGRESSION DER SIMULATORSTUDIE

Variable (* $p < 0,05$ )	Koeffizient	Exp (Koeff.)	SF	t	P
<b>Intercept (<math>\beta_0</math>)*</b>	1,54	4,76	0,10	15,25	< 0,01
<b>size (<math>\beta_1</math>)*</b>	0,35	1,42	0,09	3,78	< 0,01
<b>contrast (<math>\beta_2</math>)*</b>	0,17	1,19	0,09	2,05	< 0,05
<b>speed1 (<math>\beta_{3.1}</math>)*</b>	0,14	1,14	0,05	2,73	< 0,01
<b>speed2 (<math>\beta_{3.2}</math>)</b>	0,07	1,07	0,08	0,84	0,20

Variable (*p< 0,05)	Koeffizient	Exp (Koeff.)	SF	t	P
zss1 ( $\beta_{4.1}$ )	-0,07	0,95	0,05	-1,11	0,14
zss2 ( $\beta_{4.2}$ )*	-0,19	0,82	0,07	-2,87	< 0,05

Die Objektgröße weist eine signifikant positive Beziehung zur Reaktionszeit auf ( $p < 0,01$ ). Auch der Kontrast steht in einer signifikant positiven Beziehung zur abhängigen Variable ( $p < 0,05$ ). Die Hypothesen für die Variablen Größe und Kontrast wurden bestätigt (H1 und H2).

Es wurden statistisch signifikante Unterschiede in Bezug auf die Geschwindigkeit zwischen der Bedingung mit 40 km/h und den beiden schnelleren Bedingungen festgestellt ( $p < 0,01$ ). Die Hypothese, dass die Reaktionszeit bei langsamer Fahrgeschwindigkeit (40 km/h) höher ist als bei hoher Fahrgeschwindigkeit, kann angenommen werden. Die Beziehung zwischen den Geschwindigkeitsstufen 100 km/h und 160 km/h ist statistisch nicht signifikant ( $p = 0,20$ ). Ein möglicher Grund dafür könnte das Experimentaldesign sein, da in der Bedingung bei 160 km/h keine großen Würfel gezeigt werden konnten. Die Post-hoc-Analyse der Interaktion zwischen Geschwindigkeiten und Größe für kleine Stimuli (Abbildung 12) ergab jedoch für die Paare 100-160 km/h keine statistische Signifikanz ( $p = 0,49$ ). Daher kann nur die Hypothese über die Beziehung zwischen 40 km/h und höheren Geschwindigkeiten akzeptiert werden (H3.1).

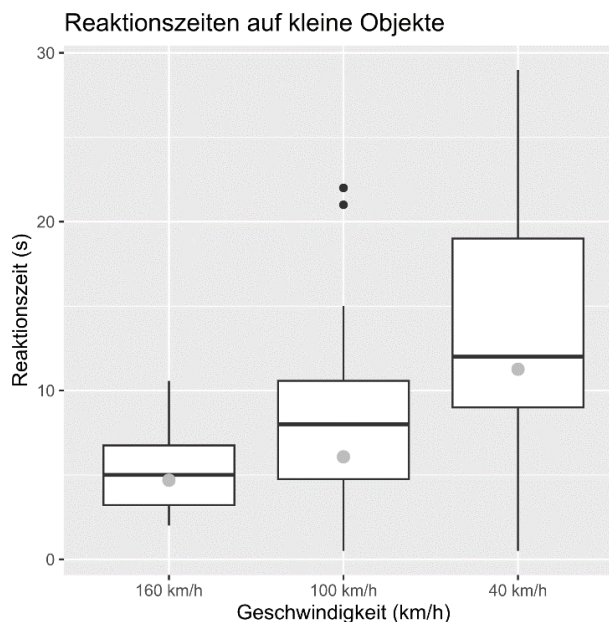


Abbildung 12: Interaktion zwischen Geschwindigkeit und Größe für kleine Stimuli (geometrische Mittelwerte als graue Punkte)

Der Vergleich zwischen dem FaS und den anderen Bedingungen erreichte keine statistische Signifikanz. Abbildung 13 zeigt den Vergleich von PZB und FaS bei der erlaubten Höchstgeschwindigkeit von 40 km/h. Unter den gleichen Bedingungen liegt das geometrische Mittel der Reaktionszeiten beim FaS-Szenario bei 8,95 s und beim PZB-Szenario bei 9,22 s. Auch der paarweise Vergleich der Szenarien PZB und FaS ergibt keine signifikanten Unterschiede. Die Analyse der Residuen zeigt jedoch, dass die Streuung der Reaktionszeitdaten unter den FaS-Bedingungen am größten ist. Es ist wichtig zu beachten, dass der Tf aufgrund der Besonderheiten des Fahrens auf Sicht bei einer zulässigen Höchstgeschwindigkeit von 40 km/h selbst über die Geschwindigkeit entscheidet. Der Variationskoeffizient dieser beiden Stichproben zeigt, dass die Variabilität der Geschwindigkeit zum Zeitpunkt der Reaktion von etwa 0,08 (8 %)

bei PZB auf etwa 0,14 (14 %) beim Fahren auf Sicht zunimmt. Dies entspricht einem Anstieg der Variabilität der Geschwindigkeit um fast 50 %.

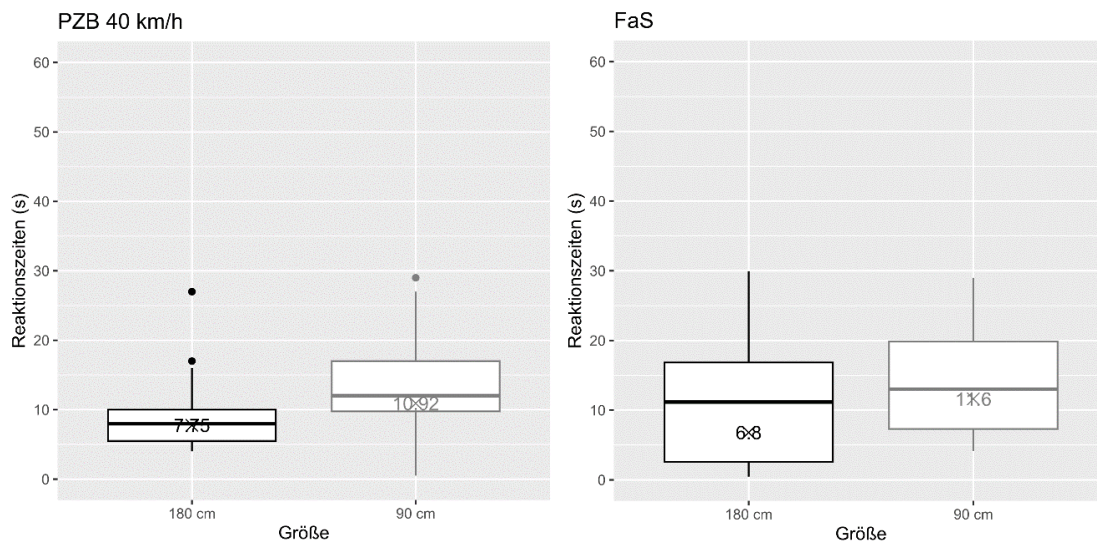


Abbildung 13: Reaktionszeiten unter PZB 40 km/h (links) und unter FaS-Bedingungen (rechts)

Für das variable Zugsicherungssystem schätzt das Modell unerwartet längere Reaktionszeiten bei PZB im Vergleich zum ETCS-System ( $\beta_{4.2} < 0$  und  $p < 0,05$ ). Eine Post-hoc-Analyse nach der Tukey-Methode zeigte jedoch, dass der signifikante Unterschied nur bei den Paaren mit unterschiedlichen Geschwindigkeitsklassen auftritt. Für den Vergleich der Zugsicherungssysteme (PZB – ETCS) bei 100 km/h erreichte der Unterschied keine Signifikanz ( $p = 0,07$ ). Dies deutet darauf hin, dass der Unterschied zwischen ETCS und PZB auf die unterschiedlichen Geschwindigkeiten zurückzuführen sein könnte und nicht auf das Zugsicherungssystem (Abbildung 14). Außerdem könnte das Bewusstsein, dass ein Objekt erscheinen wird, den Effekt der Blickverteilung bei ETCS vs. PZB abschwächen. Da das ETCS-Szenario weniger Versuchsbedingungen umfasste, war eine ausgewogene Untersuchung aller Bedingungen nicht möglich.

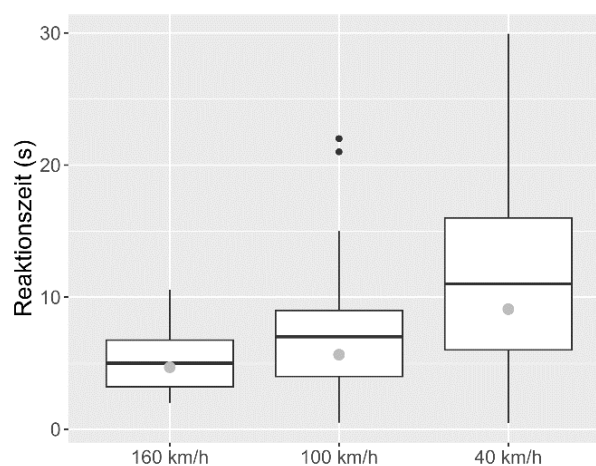


Abbildung 14: Reaktionszeit bei verschiedenen Zugsicherungssystemen (links), verschiedenen Geschwindigkeitsklassen (rechts). Die Whisker stellen 3IQR dar, die grauen Punkte sind die geometrischen Mittelwerte. Reaktionszeit-Ausreißer wurden bei der Analyse nicht ausgeschlossen.

## 8.2.1 Interpretation der Koeffizienten

Wenn man die potenzierten Koeffizienten in die Formel des Regressionsmodells (G3) einsetzt, ergibt sich die folgende Formel:

$$\log(\text{Reaktionszeit}) = 1,54 + 0,35 \cdot \text{size} + 0,17 \cdot \text{contrast} + 0,14 \cdot \text{speed}_1 + 0,07 \cdot \text{speed}_2 - 0,07 \cdot \text{zss}_1 - 0,19 \cdot \text{zss}_2 \quad (\text{G3})$$

Die potenzierten Koeffizienten entsprechen der multiplikativen Veränderung der abhängigen Variablen für eine Einheit Veränderung der entsprechenden unabhängigen Variable. Um diese Analyse durchzuführen, verwendet man das Verhältnis der geometrischen Mittelwerte für verschiedene Bedingungen. Für die Zweistufenvariablen werden diese Verhältnisse in eine prozentuale Veränderung des geometrischen Mittelwerts einer Bedingung im Vergleich zur anderen übersetzt. Da bei den Dreistufenvariablen die Stufen rezentriert wurden, stellen die potenzierten Koeffizienten prozentuale Veränderungen gegenüber dem Mittelwert auf beiden Seiten eines Vergleichs dar. Um die prozentuale Veränderung zwischen den beiden verglichenen Bedingungen zu erfassen, können die ermittelten Verhältnisse verdoppelt werden.

- Bei kleinen Reizen wird erwartet, dass sich die abhängige Variable um etwa 42 % erhöht.
- Die Verringerung des Kontrastniveaus entspricht einer 19 % Erhöhung der Reaktionszeit.
- Die Reaktionszeit bei der Geschwindigkeit von 40 km/h im Vergleich zu höheren Geschwindigkeiten ist um 28 % länger.
- Das Modell schätzt, dass bei einer Geschwindigkeit von 100 km/h die Reaktionszeit um 14 % höher ist als bei der Hochgeschwindigkeitsbedingung. Diese Beziehung ist jedoch statistisch nicht signifikant.
- Die Reaktionszeiten beim Fahren auf Sicht sind geschätzt um 10 % höher als beim Fahren mit einem Zugsicherungssystem. Jedoch erreichte dieser Vergleich keine statistische Signifikanz.
- Die vorhergesagte Reaktionszeit bei einer Fahrt mit PZB ist um 14 % länger als bei einer Fahrt mit ETCS. Dabei ist jedoch zu berücksichtigen, dass der Effekt zwischen ETCS und PZB nur bei Paaren mit unterschiedlichen Geschwindigkeitsklassen signifikant war.

Abbildung 15 zeigt die Mittelwerte der Rohreaktionszeitdaten und der logarithmierten transformierten Daten sowohl für die Stichprobe als auch für die Modellvorhersage.

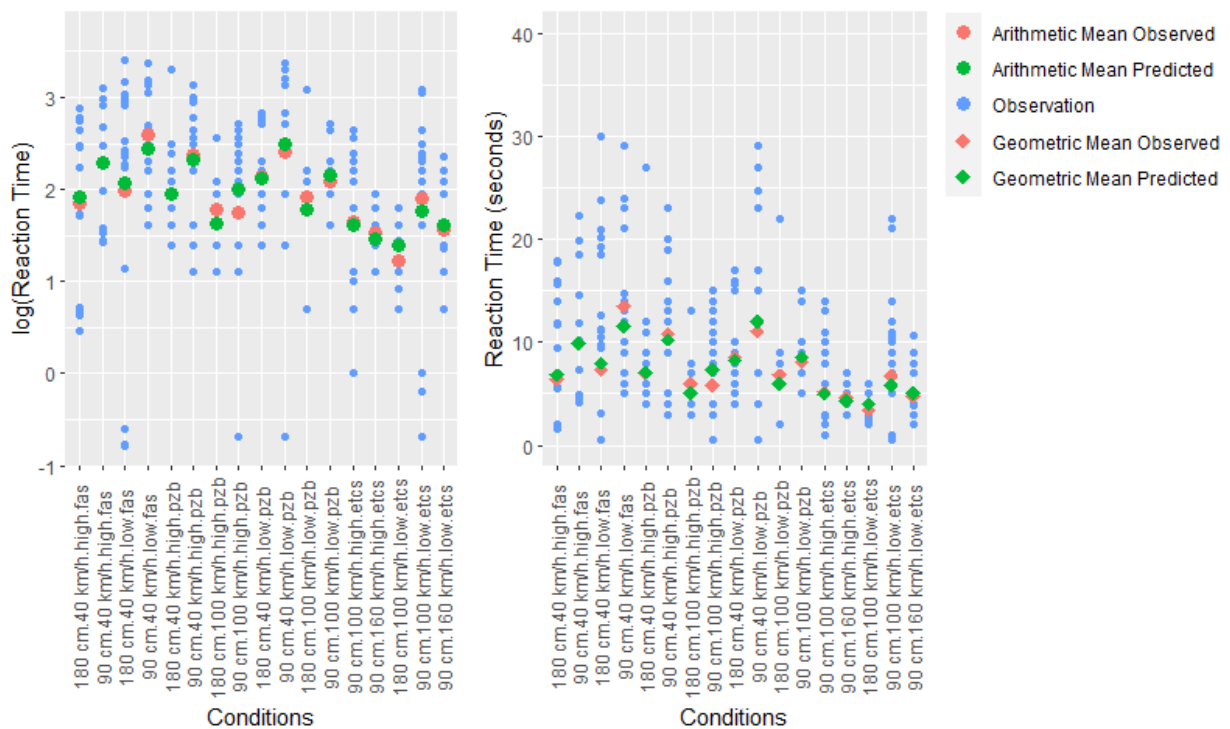


Abbildung 15: Vergleich der Modellvorhersage mit dem entsprechenden Indikator für die zentrale Tendenz unter verschiedenen Bedingungen (BBI-Studie). Log-transformierte Reaktionszeiten mit Stichprobe und vorhergesagten arithmetischen Mittelwerten (links) und rohen Reaktionszeiten mit Probe und vorhergesagten geometrischen Mittelwerten (rechts).

Um die Auswirkungen der Variablen auf die Reaktionszeit in einem ausgewogenen Design zu untersuchen, wurde noch eine multiple Regressionsanalyse nur für die PZB-Daten durchgeführt. Die Koeffizienten und die p-Werte zeigen signifikante Unterschiede zwischen den Bedingungen für alle unabhängigen Variablen im PZB-Szenario.

## 8.3 Müdigkeit, Berufserfahrung, Alter, PVT

Die Müdigkeitsbewertungen wurden vor dem Experiment und nach jeder Fahrt mithilfe der Karolinska-Schläfrigkeitsskala (KSS) erfasst. Die Skalenspunkte reichen von 1 – sehr wach bis 10 – sehr schläfrig, kann nicht wach bleiben. Block 1 und Block 2 waren die PZB- und ETCS-Abschnitte in zufälliger Reihenfolge, Block 3 war immer das Fahren auf Sicht. Die Verteilung der Antworten ist in Abbildung 16, Mitte, dargestellt, wobei die durchschnittliche Schläfrigkeit auf ähnlichem Niveau über die experimentellen Blöcke hinweg bleibt. Um die Wachsamkeit der Teilnehmenden vor jeder Fahrt zu messen, wurde ein psychomotorischer Wachsamkeitstest (PVT) durchgeführt. Mit einem Tablet wurden die Reaktionszeiten der Teilnehmenden während eines 3-minütigen PVTs aufgezeichnet. Wie bei den selbstberichteten Schläfrigkeitseinschätzungen zeigte der PVT kein Muster hinsichtlich des Zeitverlaufs der Wachsamkeit (Abbildung 16, links). Die mentale Beanspruchung der Teilnehmenden wurde nach jeder Fahrt mit dem Selbstbewertungsinstrument NASA-TLX (siehe 6.3.1) gemessen. Die durchschnittliche Arbeitsbelastung lag in allen Szenarien zwischen 41,9 und 44,9 und war bei ETCS am höchsten (Abbildung 16, rechts). Die Unterschiede in der Beanspruchung zwischen den Szenarien waren jedoch statistisch nicht signifikant ( $p > 0,05$ ).

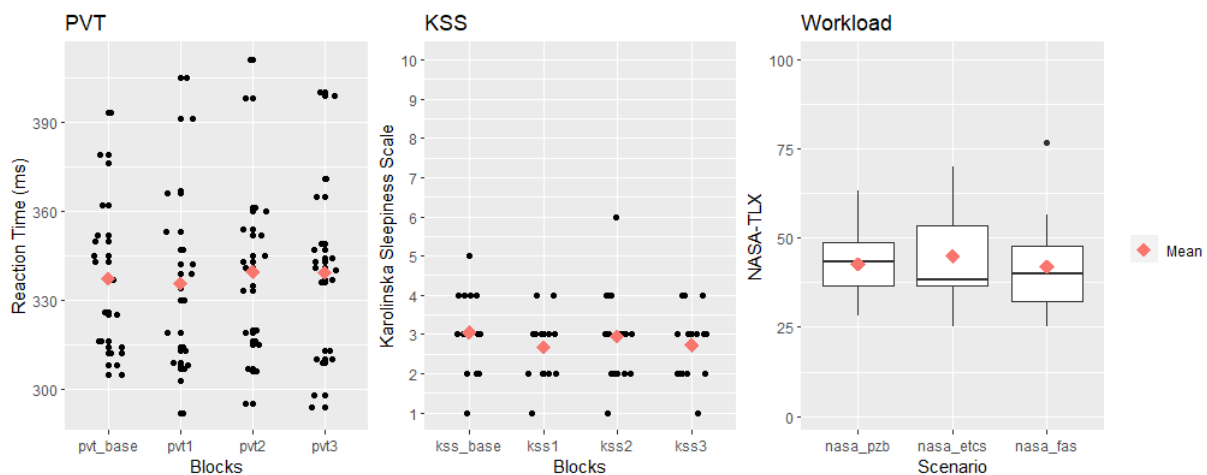


Abbildung 16: Psychomotor-Vigilanz-Test (links) und Schläfrigkeitsbewertung (Mitte) über die Versuchsblöcke der BBI-Studie. Block 1 und Block 2 waren die PZB- und ETCS-Abschnitte in randomisierter Reihenfolge, während Block 3 immer das Fahren auf Sicht war. Bewertung der Arbeitsbelastung nach jeder Fahrt (rechts). Die Durchschnittswerte sind rot markiert.

Die Reaktionszeit (PVT) korreliert positiv mit dem Alter und der Dauer der Berufserfahrung, allerdings mit einer schwachen Stärke (Spearman's Rho = 0,14 und 0,11). Die p-Werte von 0,143 und 0,064 weisen jedoch darauf hin, dass diese Korrelationen auf dem Signifikanzniveau von 0,05 statistisch nicht signifikant sind. Die Spearman-Rangkorrelation zeigt eine negative Korrelation (Rho = -0,142) zwischen der Reaktionszeit und der Vertrautheit mit dem PZB-System. Der p-Wert von 0,102 deutet jedoch darauf hin, dass diese Korrelation statistisch nicht signifikant ist. Die sehr schwache Korrelation zwischen der Reaktionszeit und der Erfahrung mit dem ETCS-System ist ebenfalls statistisch nicht signifikant (Rho = 0,067, p = 0,536).

## 8.4 Zusammenfassung der Ergebnisse

Ziel der Simulatorstudie war es, die Wahrnehmungsleistung von Tf für Objekte im Gleis oder direkt neben den Gleisen unter verschiedenen Fahrbedingungen zu messen. Die meisten der aufgestellten Forschungshypothesen wurden durch die Ergebnisse der BBI-Studie bestätigt. Die Ergebnisse zeigen, dass die Objektgröße und der Kontrast eine signifikante negative Beziehung zur Reaktionszeit aufweisen. Es wurde beobachtet, dass niedrigere Geschwindigkeiten mit längeren Reaktionszeiten verbunden sind. Die Reaktionszeit bei 40 km/h war signifikant länger als die Reaktionszeit bei höheren Geschwindigkeiten. Der Unterschied zwischen 100 km/h und 160 km/h war jedoch statistisch nicht signifikant. Die Hypothese, dass das Fahren mit ETCS im Vergleich zum Fahren mit PZB zu höheren Reaktionszeiten führt, konnte nicht bestätigt werden. Mögliche Gründe dafür könnten das unterschiedliche Geschwindigkeitsniveau und die unterschiedliche Streckencharakteristik von PZB und ETCS sein. Die Anzahl der Fahrbedingungen war für die PZB- und die ETCS-Strecke nicht gleich, deshalb war eine ausgewogene Vergleichbarkeit nicht möglich. Es ist auch möglich, dass die geringere Berufserfahrung mit dem Fahren unter ETCS zu einer erhöhten Aufmerksamkeit geführt hat. Im Gegensatz zu den Erwartungen führte das Fahren auf Sicht insgesamt nicht zu einer geringeren Reaktionszeit. Das Fahren auf Sicht führte nur bei drei der vier Bedingungen mit einer Geschwindigkeit von 40 km/h zu einer kürzeren Reaktionszeit.

## 8.5 Semistrukturiertes Interview

Nach den Experimenten wurde ein kurzes teilstrukturiertes Interview mit den Teilnehmenden (N=17) zur Wahrnehmung von Risiken und Entscheidungsfindung durchgeführt. Die Teilnehmenden bewerteten die Einschätzung des Risikos in Bezug auf verschiedene Arten von Hindernissen (von 1 - gering bis 4 - katastrophal) sowie die Häufigkeit solcher Ereignisse. Beispiele hierfür waren kleine Objekte wie Äste und Steine, große Objekte wie Bäume, verlorene Ladung und PKW, kleine und große Tiere sowie die Anwesenheit von arbeitenden oder nicht arbeitenden Personen. Zusätzlich gaben sie an, welche Risikokontrollmaßnahmen sie in solchen Situationen ergreifen.

Statische und bewegliche große Objekte werden wie erwartet als das größte Risiko wahrgenommen, mit durchschnittliche Bewertungen von 3,6 und 3,7 respektive. Allerdings traten diese Ereignisse vergleichsweise selten auf (80 % weniger als einmal im Jahr), im Vergleich zu kleineren Objekten oder Tieren auf oder in der Nähe der Gleise (Abbildung 17). Da das wahrgenommene Risiko bei großen Objekten hoch ist, umfassen die Hauptmaßnahmen die Schnellbremsung und Nothaltauftrag und ggf. Flucht aus dem Führerraum und Fdl verständigen.

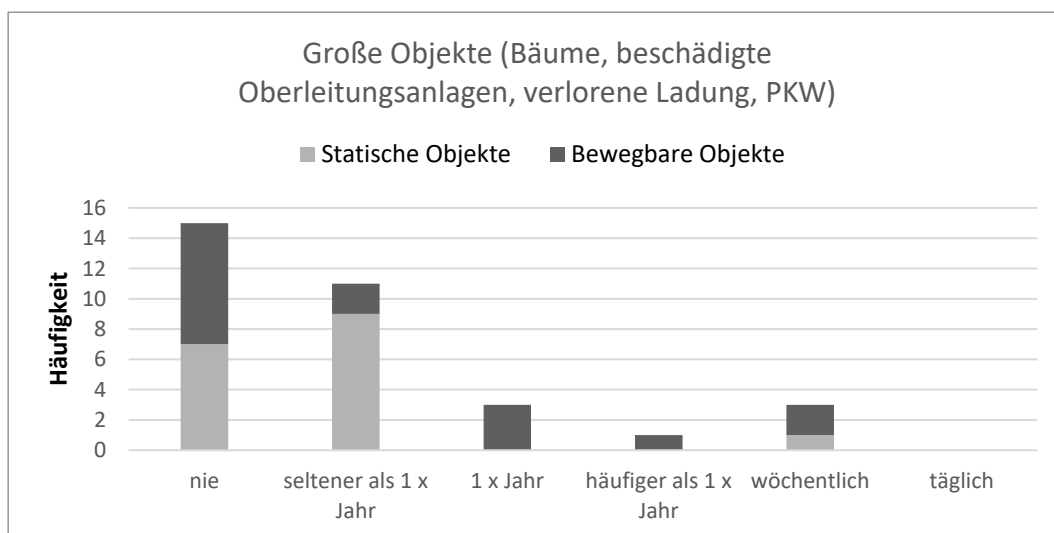


Abbildung 17: Häufigkeit des Ereignisses: Große Objekte auf dem Lichtraumprofil (N=17)

Die Häufigkeit des Auftretens von kleinen und großen Tieren (d. h. größer als ein Reh) auf oder in der Nähe der Gleise war höher, jedoch waren die Antworten hier vielfältiger. 9 von 17 Probanden berichteten von der wöchentlichen Anwesenheit kleiner Tiere, während dies bei großen Tieren bei 3 von 17 der Fall war. Die häufigsten Maßnahmen bei Anwesenheit von Tieren waren das Signalhorn betätigen, Fernlicht auf- und abblenden und Geschwindigkeit verringern. Für große Tiere erfolgte in der Regel eine schnelle Bremsung (insbesondere beim Betreten des Gleisbereichs).

Die Anwesenheit von nicht arbeitenden Personen in der Nähe der Gleise ist vergleichsweise selten (Abbildung 18), stellt jedoch ein größeres Risiko im Vergleich zu Arbeitern (durchschnittl. 3,2 und 2,5) dar. Die Häufigkeit gefährlicher Interaktionen mit den Mitarbeitern war bei den Probanden sehr unterschiedlich, was möglicherweise auf unterschiedliche Tätigkeitsbereiche zurückzuführen ist. Die Hauptmaßnahmen bei Anwesenheit von Mitarbeitern bestehen darin, zuerst die Aufmerksamkeit zu überprüfen (z. B. durch Versuch des Blickkontakts) und bei Bedarf die Geschwindigkeit zu reduzieren oder eine

schnelle Bremsung einzuleiten. Allerdings werden konservativere Maßnahmen ergriffen, wenn sich Personen in der Nähe oder auf den Gleisen befinden und keine Arbeiter sind. Hierzu gehören das Betätigen der Signalhupe, Schnellbremsung, mindestens eine Geschwindigkeitsreduktion und die Fdl über die Person auf den Gleisen informieren.

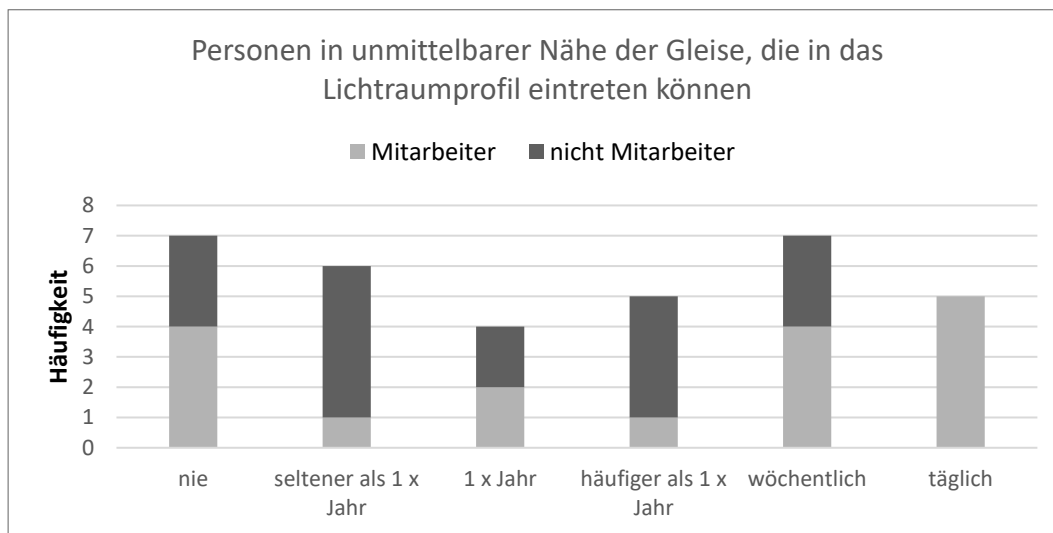


Abbildung 18: Häufigkeit des Ereignisses: Personen auf dem Lichtraumprofil (N=17)

Die Teilnehmenden haben auch die Wichtigkeit verschiedener Faktoren für ihre Entscheidungsfindung bewertet. Außerdem haben sie ihre Bewertung schriftlich oder mündlich begründet. Der Objektart, die Größe der Hindernisse, der Abstand zu den Gleisen und die Geschwindigkeit des Zuges wurden mit durchschnittlichen Bewertungsraten von 4,3 bis 4,9 als sehr wichtig bewertet. Das geschätzte Schadensausmaß war die Hauptüberlegung bei der Entscheidungsfindung bei der Bewertung dieser Faktoren. Es wurden auch folgende Faktoren genannt, die die Entscheidungsfindung für Risikokontrollmaßnahmen beeinflussen: das Gewicht des Zuges, das Bremsverhalten des Zuges, die Zugklasse, das Wetter, die Tageszeit, die Landschaft. Ein Teilnehmender erwähnte den potenziellen Ablenkungseffekt beim Fahren mit Buchfahrplänen (z. B. La), bei dem Tf möglicherweise die korrekten Informationen finden muss.

## 9 Ergebnisse Simulatorstudie DLR

### 9.1 Deskriptive Analyse der Reaktionszeitdaten

In Abbildung 19 werden die erhobenen Daten für die verschiedenen Experimentalbedingungen mit ihrem arithmetischem Mittel, dem geometrischen Mittel sowie dem Median dargestellt. In allen Bedingungen waren der Median und das geometrische Mittel ähnlich, während das arithmetische Mittel höhere Werte aufwies. Die Gesamtverteilung der Reaktionszeiten ist in Abbildung 20 dargestellt und bestätigt die typische nicht-normale Verteilung von Reaktionszeitmessungen mit einer unteren Grenze knapp über 0 Sekunden und einem langen Verteilungsende auf der rechten Seite. In Anbetracht dieser Verteilung sind der Median und das geometrische Mittel als Maß für die zentrale Tendenz dem arithmetischen Mittel für die Reaktionszeit vorzuziehen, da sie robuster gegenüber Ausreißern sind. Die geometrischen Mittelwerte mit ihren Standardabweichungen sind für die einzelnen Bedingungen in Tabelle 14 dargestellt

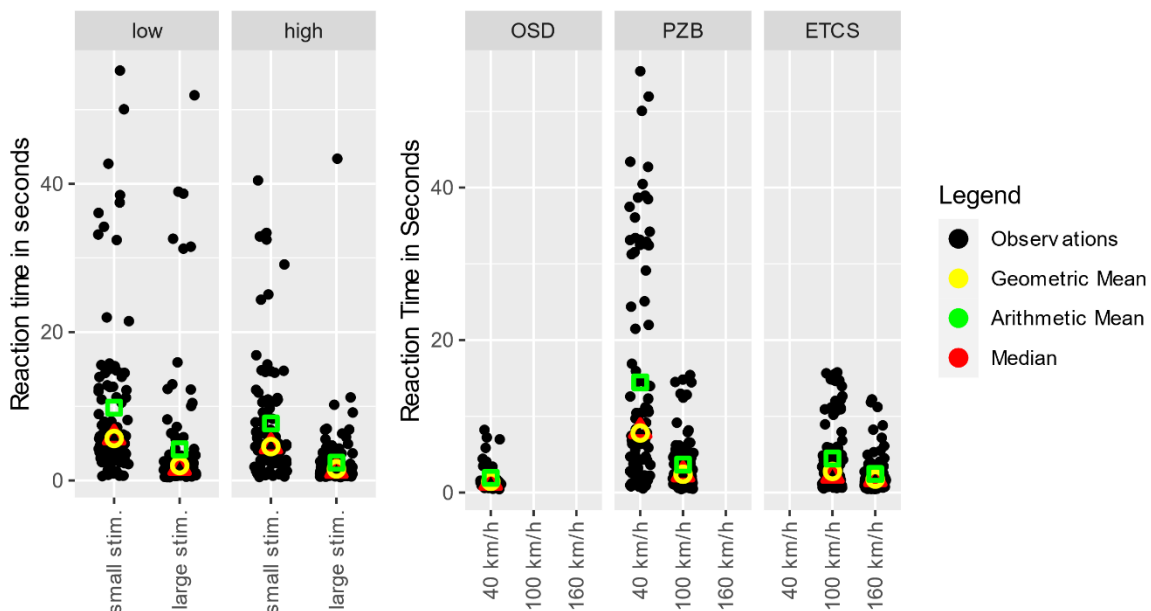


Abbildung 19: Maße der zentralen Tendenz für einige Stimulus-Eigenschaftsbedingungen (links) und Zugbetriebsparameterbedingungen (rechts) in der Simulatorstudie. Geometrische Mittelwerte sind gelb dargestellt, arithmetische Mittelwerte grün und Medianwerte rot.

Betrachtet man die geometrischen Mittelwerte und Mediane, so zeigen die Daten wie erwartet längere Reaktionszeiten für kleine Stimuli im Vergleich zu großen Stimuli sowie für kontrastarme Stimuli im Vergleich zu kontrastreichen Stimuli. Zudem reagierten die Teilnehmenden wie erwartet bei höheren Geschwindigkeiten schneller als bei niedrigeren Geschwindigkeiten. Entgegen der Erwartungen waren jedoch die Reaktionszeiten in der PZB-Bedingung länger als in der ETCS-Bedingung. Tabelle 14 zeigt eine Übersicht der gemessenen Mittelwerte und Standardabweichungen pro Variablenkombination.

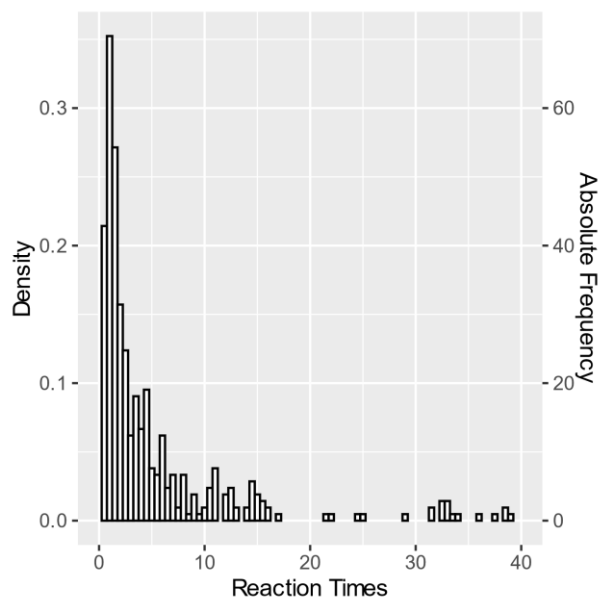


Abbildung 20: Histogramm der in der DLR-Simulatorstudie gemessenen Reaktionszeiten

TABELLE 14: MITTELWERTE UND STANDARDABWEICHUNGEN DER REAKTIONSZEIT DER VERSCHIEDENEN VARIABLENKOMBINATIONEN IN DER DLR-SIMULATORSTUDIE.

	40 km/h			100 km/h			160 km/h		
<b>FaS</b>	kontrast-reich	groß	1,42 (1,94)						
	kontrast-arm	groß	1,51 (2,07)						
<b>PZB</b>	kontrast-reich	groß	2,90 (2,75)	kontrast-reich	groß	1,33 (1,86)			
		klein	12,68 (2,33)		klein	3,18 (2,14)			
	kontrast-arm	groß	6,42 (3,97)	kontrast-arm	groß	1,56 (2,33)			
		klein	15,73 (2,53)		klein	4,95 (2,48)			
<b>ETCS</b>				kontrast-reich	groß	1,45 (2,05)	kontrast-reich	groß	1,18 (1,83)
					klein	4,81 (2,60)		klein	2,31 (2,51)
				kontrast-arm	groß	1,50 (1,64)	kontrast-arm	groß	1,53 (1,70)
					klein	5,94 (2,17)		klein	2,32 (2,33)

## 9.2 Ergebnisse der Regressionsanalyse

Die Auswirkungen der unabhängigen Variablen auf die Reaktionszeit wurden mithilfe einer Regressionsanalyse analysiert. Es wurde die gleiche Modellierung wie für Tabelle 13 verwendet. Die Ergebnisse, einschließlich der Koeffizienten des Modells und die Teststatistiken, sind in Tabelle 15 zu finden. Das Modell erklärte knapp 50 % der Varianz in der Stichprobe ( $R^2 = 0,4535$ ,  $R^2_{adj} = 0,446$ ). Die Koeffizienten in der Tabelle entsprechen den Parametern des linearen Modells – z. B. entspricht  $b_1$  dem Modellparameter  $\beta_1$  (die Auswirkung der Stimulusgröße auf das Modell). Da die Daten in eine logarithmische Skala umgewandelt wurden, wurden auch exponentielle Koeffizienten berechnet. Die  $p$ -Werte in der Tabelle geben das Signifikanzniveau für den einseitigen Test der Koeffizienten gegen Null an. Das heißt, der  $p$ -Wert beschreibt die Wahrscheinlichkeit, ein Ergebnis zu erhalten, das genauso groß oder größer ist als das beobachtete Ergebnis, wenn man annimmt, dass der wahre Parameter gleich Null ist. Wenn  $p < 0,05$ , kann die Hypothese, dass der wahre Parameter größer als Null ist, angenommen werden. Alle in Kapitel 7 formulierten Hypothesen gehen davon aus, dass ein Parameter größer als Null ist (vgl. H1, H2, H3, H4), dementsprechend können die Hypothesen angenommen werden, wenn der entsprechende  $p$ -Wert  $< 0,05$  ist.

TABELLE 15: REGRESSIONSKOEFFIZIENTEN UND T-WERTE FÜR DIE LINEARE REGRESSION DER DLR-SIMULATOR-STUDIE. MULTIPLE R-SQUARED: 0.4559, ADJUSTED R-SQUARED: 0.4481.

Variable	Koeffizient	Potenzierter Koeffizient	Standardfehler	t	p	r
<b>a - Intercept</b>	1.39	4.03	0.09	15.44	< .001	0.61
<b>b1 - size</b>	1	2.72	0.09	11.65	< .001	0.50
<b>B2 - contrast</b>	0.24	1.27	0.08	3.01	< .01	0.15
<b>B3.1 - speed</b>	0.47	1.60	0.05	10.32	< .001	0.45
<b>B3.2 - speed2</b>	0.23	1.26	0.06	3.82	< .001	0.18
<b>B4.1 - zss1</b>	0.42	1.52	0.06	7.75	< .001	0.35
<b>B4.2 - zss2</b>	0.07	1.07	0.06	1.28	0.20	0.06

Wie aus Tabelle 15 hervorgeht, hatte die Stimulusgröße einen signifikanten Einfluss auf die Reaktionszeit, mit  $p < 0,01$ . Daher kann H1: „Die Reaktionszeit ist bei kleinen Stimuli höher.“ angenommen werden. Der Kontrast hatte ebenfalls einen signifikanten Einfluss auf die Reaktionszeit ( $p < 0,01$ ). Daher kann auch H2: „Die Reaktionszeit ist bei kontrastarmen Reizen höher.“ angenommen werden.

In Bezug auf die Geschwindigkeit wurden signifikante Unterschiede zwischen 40 km/h und den beiden schnelleren Bedingungen identifiziert ( $p < 0,01$ ). Auch zwischen den Bedingungen 100 km/h und 160 km/h lag ein signifikanter Unterschied vor ( $p < 0,01$ ). Daher können H3a: „Die Reaktionszeit ist bei langsamer Fahrgeschwindigkeit (40 km/h) höher als bei den Bedingungen mit hoher Fahrgeschwindigkeit.“ sowie H3b: „Die Reaktionszeit ist bei einer Fahrgeschwindigkeit von 100 km/h höher als bei 160 km/h.“ angenommen werden.

Wie bei der Geschwindigkeit zeigte sich ein signifikanter Effekt auf die Reaktionszeit beim Fahren auf Sicht im Vergleich zum Fahren mit einem Zugsicherungssystem (H4a: „Die Reaktionszeit ist höher, wenn ein Zugsicherungssystem verwendet wird, als beim Fahren auf Sicht.“;  $p < 0,01$ ). Die Hypothese kann

folglich angenommen werden. Anders als erwartet wurde jedoch kein signifikanter Unterschied zwischen den Zugsicherungssystemen PZB und ETCS gefunden. Hypothese H4b: „Die Reaktionszeit ist höher, wenn mit ETCS gefahren wird, als mit PZB.“ kann demnach nicht angenommen werden.

## 9.2.1 Interpretation der Koeffizienten

Da die Reaktionszeitdaten vor der Regressionsanalyse log-transformiert wurden, können die Koeffizienten des Modells nicht ohne weiteres direkt interpretiert werden. Allerdings können unter bestimmten Voraussetzungen die potenzierten Regressionskoeffizienten ( $\exp(\beta)$ ) zur Interpretation herangezogen werden. Zunächst zieht man hierfür das Verhältnis der geometrischen Mittelwerte unter den verschiedenen Bedingungen heran. Für die mit Dummies kodierten 2-Ebenen-Faktoren werden diese Verhältnisse dann in eine prozentuale Veränderung des geometrischen Mittels in einer Bedingung gegenüber der anderen übersetzt. Wenn zum Beispiel ein potenziertes Koeffizient für den Faktor Größe 1,4 beträgt, würde dies bedeuten, dass die Reaktionszeit für kleine Stimuli um 40 % höher war als die Reaktionszeit für große Stimuli.

Bei den Dreistufenvariablen stellen die potenzierten Koeffizienten prozentuale Veränderungen vom Mittelwert beider Seiten eines Vergleichs dar. Diese können somit verdoppelt werden, um die prozentuale Veränderung zwischen den beiden verglichenen Bedingungen zu erfassen. Für den Vergleich zwischen der Bedingung 40 km/h und dem Mittelwert der beiden schnelleren Geschwindigkeiten würde dies bei einem Koeffizienten von 1,4 bedeuten, dass die Reaktionszeit in der langsamen Bedingung 80 % höher war als unter den schnelleren Bedingungen.

In Tabelle 15 sind die potenzierten Koeffizienten des Modells angegeben. Setzt man die Koeffizienten in die Formel des Regressionsmodells ein, erhält man die folgende Formel:

$$\begin{aligned} \text{Reaktionszeit} = & \exp(4.03) \cdot \exp(2.72)^{\text{size}} \cdot \exp(1.27)^{\text{contrast}} \cdot \exp(1.6)^{\text{speed1}} \cdot \exp(1.26)^{\text{speed2}} \\ & \cdot \exp(1.52)^{\text{zss1}} \cdot \exp(1.07)^{\text{zss2}} \end{aligned} \quad (\text{G4})$$

Anhand der Interpretation der potenzierten Koeffizienten lassen sich die folgenden Aussagen ableiten:

- Bei kleinen Stimuli ist die vorhergesagte Reaktionszeit 172 % länger als bei großen Stimuli.
- Bei einem geringeren Kontrast ist die vorhergesagte Reaktionszeit um 27 % länger als bei einem höheren Kontrast.
- Bei 40 km/h ist die vorhergesagte Reaktionszeit 80 % länger als bei höheren Geschwindigkeiten.
- Bei 100 km/h ist die vorhergesagte Reaktionszeit 52 % länger als bei einer Geschwindigkeit von 160 km/h.
- Wird mit PZB oder ETCS gefahren, ist die vorhergesagte Reaktionszeit 104 % länger als beim Fahren auf Sicht.
- Bei einer Fahrt mit PZB ist die vorhergesagte Reaktionszeit 14 % länger als bei einer Fahrt mit ETCS. Dabei ist jedoch zu berücksichtigen, dass der Effekt zwischen ETCS und PZB nicht signifikant war.

## 9.3 Einfluss von Müdigkeit, Berufserfahrung, Alter

Die Daten zu Geschlecht, Alter und Berufserfahrung wurden gesammelt und basierend auf Korrelationen ausgewertet. Da alle Teilnehmenden an der Simulatorstudie männlich waren, konnte der Einfluss des Geschlechts auf die Reaktionszeit in dieser Studie nicht bewertet werden. Das Alter hatte eine sehr

minimale Korrelation mit der Reaktionszeit (Spearman's Rank Korrelation,  $\rho = 0,062$ ), die Korrelation für die Berufserfahrung war ebenfalls sehr gering (Spearman's Rank Korrelation,  $\rho = 0,108$ ). Die Vertrautheit mit den verschiedenen Zugsicherungssystemen korrelierte nur schwach mit den Reaktionszeiten bei der Nutzung der jeweiligen Zugsicherungssysteme, sowohl für PZB (Spearman's Rank Korrelation,  $\rho = -0,005$ ), als auch für ETCS (Spearman's Rank Korrelation,  $\rho = -0,031$ ).

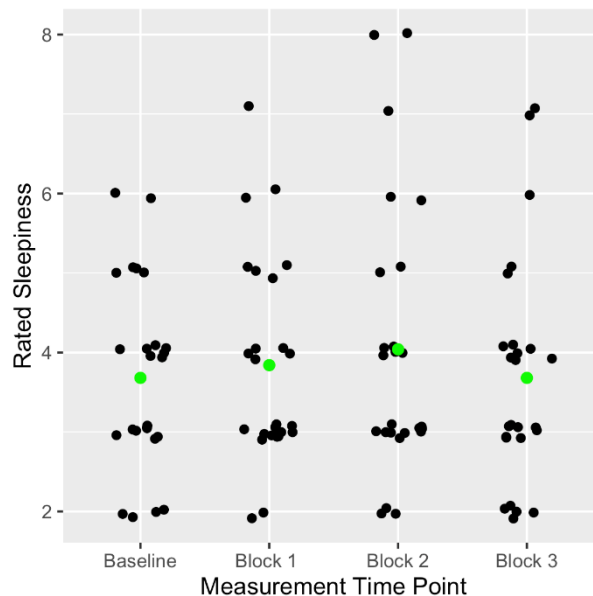


Abbildung 21: Schläfrigkeitsbewertung über die Versuchsblöcke der DLR-Simulatorstudie. Die Durchschnittswerte sind grün markiert. Block 1 und Block 2 waren die PZB- und ETCS-Abschnitte in randomisierter Reihenfolge, während Block 3 immer das Fahren auf Sicht war.

Die Müdigkeit wurde mit der Karolinska-Sleepiness-Scale (KSS) vor dem Experiment und nach jedem Versuchsblock erhoben. Alle Teilnehmenden bewerteten ihre Müdigkeit zwischen "sehr wach" und "einige Anzeichen von Schläfrigkeit". Im Verlauf des Experiments stieg die durchschnittliche Müdigkeit zunächst leicht an und nahm dann nach dem letzten Block ab, ohne dass ein klares Muster erkennbar ist (Abbildung 21). In Bezug auf den Einfluss der Müdigkeit auf die Reaktionszeit ließ sich kein klares Muster erkennen.

## 9.4 Zusammenfassung der Ergebnisse

Das Hauptziel der Simulatorstudie bestand darin, die Wahrnehmungsleistung von Tf für einen Gegenstand auf oder direkt neben den Gleisen zu bemessen. Die Ergebnisse der DLR-Studie bestätigen die meisten der auf früheren Untersuchungen basierenden Hypothesen. Sie zeigen, dass Stimulusgröße und -kontrast sowie die Zuggeschwindigkeit einen deutlichen Einfluss auf die Reaktionszeit haben. Die Ergebnisse stimmen auch mit der Hypothese überein, dass das Fahren auf Sicht zu kürzeren Reaktionszeiten führt, da die Aufmerksamkeit stärker auf die Gleisumgebung gerichtet ist. Die Hypothese, dass das Fahren mit ETCS zu höheren Reaktionszeiten im Vergleich zum Fahren mit PZB führt, konnte jedoch nicht bestätigt werden. Ein Grund dafür könnten die Unterschiede in der Streckengestaltung zwischen ETCS- und PZB-Strecken sein. Die PZB-Strecke ist im Hinblick auf Vegetation und Gebäude abwechslungsreicher, die Strecke ist kurvenreicher und wird nicht von Oberleitungsmasten eingerahmt. Dies bietet möglicherweise mehr auffällige Reize außerhalb der Gleisumgebung, was zu einer stärkeren visuellen

Erkundung abseits der Gleise und damit zu höheren Reaktionszeiten führt. Außerdem zeigte eine visuelle post-hoc Auswertung von Videodaten, dass die Tf in sechs Fällen von Reaktionszeiten über drei Standardabweichungen im Moment des Stimulus-Auftauchens auf die Konsole schauten.

# 10 Ergebnisse Onlinestudie DLR

## 10.1 Deskriptive Analyse der Reaktionszeitdaten

Wie aus Abbildung 22 hervorgeht, lagen der Median und das geometrische Mittel nahe beieinander. Das arithmetische Mittel lag ähnlich der Simulatorstudie über die Bedingungen hinweg höher. Die Verteilung der Reaktionszeitdaten in Abbildung 23 zeigt ein sehr ähnliches Muster wie bei der Simulatorstudie-DLR. Tabelle 16 zeigt eine Übersicht der gemessenen Mittelwerte und Standardabweichungen.

Betrachtet man die geometrischen Mittelwerte und Mediane, so zeigen die Daten längere Reaktionszeiten für kleine Stimuli im Vergleich zu großen Stimuli sowie für kontrastarme Stimuli im Vergleich zu kontrastreichen Stimuli. Bei höheren Geschwindigkeiten reagierten die Teilnehmenden schneller als bei niedrigeren Geschwindigkeiten. Und schließlich waren die Reaktionszeiten in der PZB-Bedingung länger als in der ETCS-Bedingung (Tabelle 16).

TABELLE 16: MITTELWERTE UND STANDARDABWEICHUNGEN DER VERSCHIEDENEN VARIABLENKOMBINATIONEN IN DER ONLINE-STUDIE.

	40 km/h			100 km/h			160 km/h		
<b>PZB</b>	kontrast-reich	groß	1,73 (2,14)	kontrast-reich	groß	1,43 (1,78)			
		klein	3,89 (2,41)		klein	4,12 (1,79)			
	Kontrast-arm	groß	4,73 (1,90)	kontrast-arm	groß	3,57 (1,72)			
		klein	7,70 (1,73)		klein	4,81 (2,11)			
<b>ETCS</b>				kontrast-reich	groß	1,67 (1,99)	kontrast-reich	groß	1,28 (1,52)
					klein	7,18 (1,59)		klein	4,53 (1,82)
				kontrast-arm	groß	1,81 (1,96)	kontrast-arm	groß	1,27 (2,08)
					klein	1,48 (1,93)		klein	2,21 (1,86)

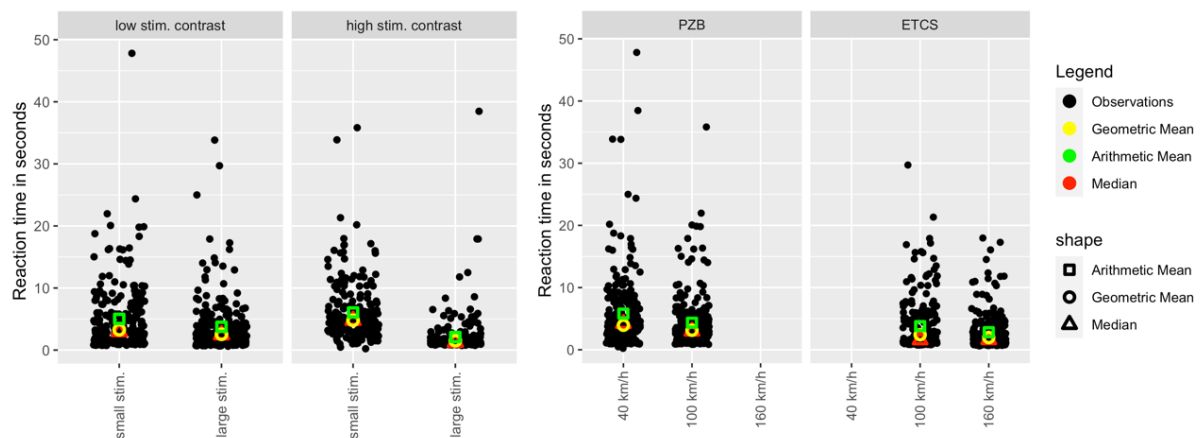


Abbildung 22: Maße der zentralen Tendenz für einige Stimulus Eigenschaftsbedingungen (links) und Zugbetriebsparameterbedingungen (rechts) in der DLR-Simulatorstudie. Geometrische Mittelwerte sind gelb dargestellt, arithmetische Mittel grün und Medianwerte rot.

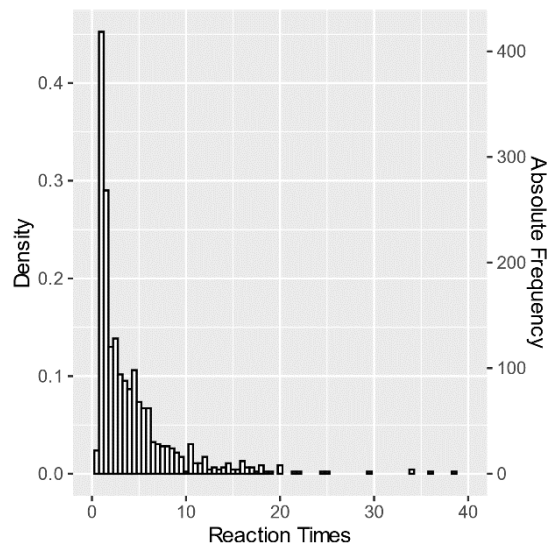


Abbildung 23: Diagramm der in der Online-Studie gemessenen Reaktionszeiten.

## 10.2 Ergebnisse der Regressionsanalyse

Das Modell für die Onlinestudie klärte etwa 25 % der Varianz in der Stichprobe auf, also deutlich weniger als bei der Simulatorstudie ( $R^2 = 0,25$ ,  $R^2_{adj} = 0,2459$ ). Ein Überblick über die Ergebnisse der Regressionsanalyse für die Onlinestudie einschließlich der Koeffizienten des Modells und der t-Test-Statistiken findet sich in Tabelle 17. Die Hypothesen wurden angenommen, wenn der entsprechende p-Wert des Tests gleich oder kleiner als 0,05 war.

TABELLE 17: REGRESSIONSKOEFFIZIENTEN UND T-WERTE FÜR DIE LINEARE REGRESSION DER ONLINESTUDIE.

Variable	Koeffizient	Potenzierter Koeffizient	Standard fehler	t	p	r
<b>a - Intercept</b>	0.51	1.66	0.07	7.45	< .01	0.24
<b>b1 - size</b>	0.71	2.03	0.04	16.1	< .01	0.47
<b>B2 - contrast</b>	0.09	1.09	0.04	1.94	0.03	0.06
<b>B3.1 - speed</b>	0.11	1.11	0.02	4.76	< .01	0.15
<b>B3.2 - speed2</b>	0.1	1.11	0.03	3.28	< .01	0.11
<b>B4 - zss1</b>	0.27	1.31	0.06	4.31	< .01	0.14

Wie in der Simulatorstudie zeigte die Regressionsanalyse einen signifikanten Effekt der Stimulusgröße, des Kontrasts und der Geschwindigkeit. Die Reaktionszeit war bei kleinen Stimuli höher als bei großen Stimuli und höher bei kontrastarmen Stimuli als bei kontrastreichen. Hypothesen H1 und H2 konnten demzufolge angenommen werden ( $p < 0,05$ ). Die Reaktionszeit war bei 40 km/h höher als bei den beiden höheren Geschwindigkeiten und bei 100 km/h höher als bei 160 km/h. Auch Hypothesen H3a und H3b konnten also angenommen werden ( $p < 0,05$ ). Anders als bei der Simulatorstudie war jedoch der Unterschied zwischen PZB und ETCS signifikant ( $p < 0,01$ ). Die Probanden reagierten langsamer auf Stimuli auf der PZB-Strecke als auf Stimuli auf der ETCS-Strecke. Hypothese H4: „Die Reaktionszeit ist niedriger, wenn mit ETCS gefahren wird als mit PZB.“ kann demnach angenommen werden.

### 10.2.1 Interpretation der Koeffizienten

Da auch in der Analyse der Online-Studie log-transformierte Daten verwendet werden, können die potenzierten Koeffizienten wie in Kapitel 9.2.1 beschrieben interpretiert werden. Anhand der Interpretation der potenzierten Koeffizienten lassen sich die folgenden Aussagen ableiten:

- Bei kleinen Stimuli ist die vorhergesagte Reaktionszeit 103 % länger als bei großen Stimuli.
- Bei einem geringeren Kontrast ist die vorhergesagte Reaktionszeit um 9 % länger als bei einem höheren Kontrast.
- Bei 40 km/h ist die vorhergesagte Reaktionszeit 22 % länger als bei höheren Geschwindigkeiten.
- Bei 100 km/h ist die vorhergesagte Reaktionszeit 22 % länger als bei einer Geschwindigkeit von 160 km/h.
- Wird mit PZB gefahren, ist die vorhergesagte Reaktionszeit 31 % länger als beim Fahren auf Sicht.
- Bei einer Fahrt mit PZB ist die vorhergesagte Reaktionszeit 16 % länger als bei einer Fahrt mit ETCS.

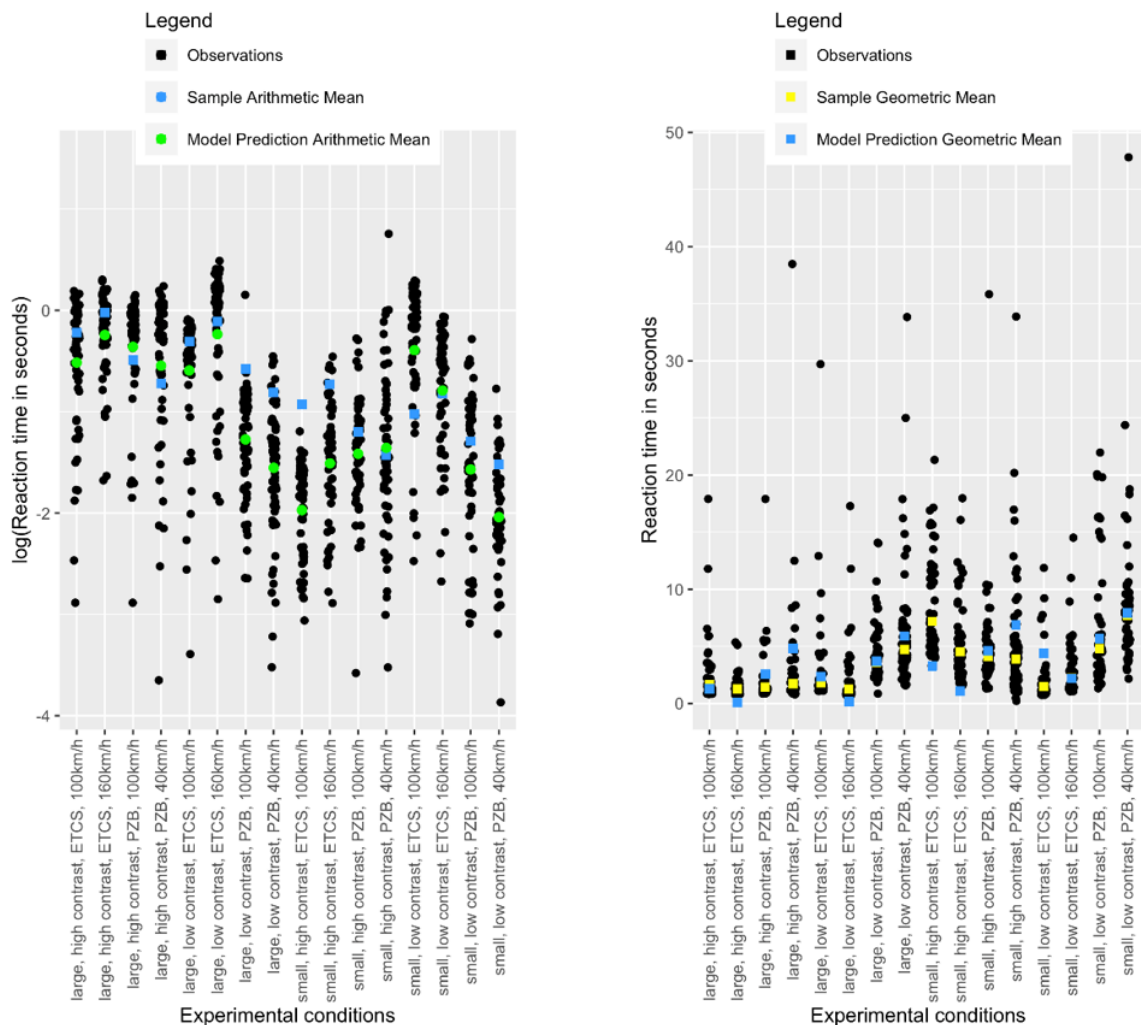


Abbildung 24: Vergleich der Modellvorhersage mit dem entsprechenden Indikator der zentralen Tendenz über die Bedingungen hinweg (Online-Studie). Logarithmisch transformierte Reaktionszeiten mit arithmetischen Mittelwerten der Stichprobe und der Vorhersage (links) und rohe Reaktionszeitdaten mit Probe und vorhergesagten geometrischen Mittelwerten (rechts).

Abbildung 24 zeigt die vorhergesagten geometrischen Mittelwerte für jede Bedingung auf der Grundlage des Regressionsmodells im Vergleich zu den beobachteten geometrischen Mittelwerten in der Stichprobe.

## 10.2.2 Einfluss von Müdigkeit, Geschlecht, Alter sowie Vorladen des Videos

Genau wie in der Simulatorstudie wurden die demografischen Daten nicht in das lineare Modell aufgenommen. Stattdessen wurden diese Daten deskriptiv ausgewertet. Alter hatte eine minimale Korrelation mit der Reaktionszeit (Spearman's Rank Korrelation,  $\rho = -0,087$ ). Vergleicht man die geometrischen Mittelwerte der Reaktionszeit für das Geschlecht, so war der Unterschied ebenfalls gering ( $GM_{\text{male}} = 2,640$  s,  $GM_{\text{female}} = 2,496$  s).

Die Bewertungen der Müdigkeit wurden in der Onlinestudie vor und nach dem Experiment mithilfe der Karolinska-Sleepiness-Scale erhoben. Die Bewertungen der Teilnehmenden variierten stärker als in der

Simulatorstudie. Die Bewertung der Müdigkeit lag bei der Erhebung vor Beginn des Experiments zwischen den Werten 1 = "extrem wach" und 8 = "schläfrig, mit einiger Anstrengung wach zu bleiben". Im Mittel wurde die Müdigkeit mit 3,72 angegeben. Im Verlauf des Experiments stieg die mittlere Schläfrigkeit auf einen Wert von 4,27 nach Beendigung des Experiments. Das Auftragen der Schläfrigkeit gegen die Reaktionszeit ergab jedoch ähnlich wie in der Simulatorstudie kein klares Muster.

Für die meisten Teilnehmenden der Stichprobe war eine Teilnahme am Experiment mit einem Vorladen (Preload) aller Videos technisch möglich. Für 19 von 70 Teilnehmenden war der Preload aufgrund technischer Kompatibilitätsprobleme nicht möglich. Die Analyse der Reaktionszeiten der beiden Gruppen ergab, dass die Reaktionszeiten in der On-Demand-Video-Gruppe ( $GM_{\text{preload}} = 2,560$  s,  $GM_{\text{on-demand}} = 3,461$  s) im Schnitt höher waren. Dies deutet darauf hin, dass die Ladezeit des On-Demand-Videos zu einer Verzögerung der Reaktionszeit der Teilnehmenden führte.

## 10.3 Zusammenfassung der Ergebnisse

Das Hauptziel der Onlinestudie bestand darin, die Ergebnisse aus der Simulatorstudie durch eine größere Stichprobe zu ergänzen. Die Ergebnisse der Studie zeigen zu großen Teilen die gleichen Ergebnisse wie die Simulatorstudien. In allen Studien wurde festgestellt, dass größere Stimuli schneller erkannt wurden. Ausgehend von der Interpretation der exponentiellen Modellkoeffizienten und den Mittelwerten hatte die Variable Größe in allen Studien den stärksten Einfluss auf die vorhergesagte Reaktionszeit. Ebenso wurde in allen Studien festgestellt, dass ein geringerer Kontrast zu signifikant längeren Reaktionszeiten führt. Dabei war jedoch der Einfluss des Kontrasts in der Onlinestudie geringer als in den Simulatorstudien. Dies könnte auf weniger kontrollierte Bildschirmaufbauten einschließlich kleinerer Bildschirmgrößen und Beleuchtungsunterschieden in der Onlinestudie zurückzuführen sein. Die Geschwindigkeit des Zugs hatte ebenfalls in allen Studien einen signifikanten Einfluss auf die Reaktionszeit. In allen Studien wurden Stimuli bei höheren Geschwindigkeiten schneller erkannt. Diese Ergebnisse könnten darauf hindeuten, dass die angenommene Vertikalisation des Blicks bei schnelleren Geschwindigkeiten nicht nur bei der Fahrt im Simulator auftritt, sondern auch bei der Betrachtung von Videos. Sicherlich hatte auch einen Effekt, dass bei höherer Geschwindigkeit die Objekte optisch schneller groß werden und daher schneller erkannt werden.

In den Simulatorstudien konnte kein signifikanter Unterschied zwischen den Reaktionszeiten bei PZB und ETCS festgestellt werden. Es wurde vermutet, dass dies eine Folge der Unterschiede in der Gleisumgebung zwischen beiden Strecken war. Im Rahmen der Onlinestudie wurden dann signifikant höhere Reaktionszeiten bei den Stimuli auf der PZB-Strecke festgestellt als bei den Stimuli auf der ETCS-Strecke, obwohl die Teilnehmenden in der Onlinestudie keine Nebenaufgaben im Zusammenhang mit dem Zugsicherungssystem erledigen mussten und ihre volle Aufmerksamkeit auf die Gleise richten konnten. Diese Ergebnisse könnten als Bestätigung der Theorie herangezogen werden, dass Unterschiede in der Gestaltung der Gleisumgebung und damit Unterschiede in der visuellen Exploration die Reaktionszeiten beeinflussen. Es wurde vermutet, dass die Reaktionszeiten auf der PZB-Strecke in den Simulatorstudien nicht signifikant kürzer waren als die Reaktionszeiten auf der ETCS-Strecke, da die PZB-Strecke im Hinblick auf Vegetation und Gebäude abwechslungsreicher und kurvenreicher ist und nicht von Oberleitungsmasten eingerahmt wird. Dies bietet möglicherweise mehr auffällige Reize außerhalb der Gleisumgebung, führt zu einer stärkeren visuellen Erkundung abseits der Gleise und damit zu höheren Reaktionszeiten.

# 11 Analyse der Gesamtstudie

Die erhobenen Daten aus zwei Simulatorstudien wurden zusammengeführt, um die Reaktionszeiten von Tf unter verschiedenen Bedingungen zu analysieren. Die Onlinestudie wurde von dieser Analyse ausgeschlossen, da sie mit Probanden durchgeführt wurde, die keine Tf waren, und dies in einer anderen Versuchsumgebung stattfand.

Dieses Kapitel präsentiert die Ergebnisse der Regressionsanalyse sowie der Analysen ohne Verwendung zentraler Tendenzen, um fundierte Einsichten in die Auswirkungen der unabhängigen Variablen auf die Reaktionszeiten zu liefern. Die Ergebnisse wurden sowohl graphisch als auch statistisch dargestellt und ermöglichten eine detaillierte Interpretation der Effekte. Darüber hinaus wurde die Überlebensanalyse verwendet, um das Verhalten der RT im Hinblick auf das Ereignis der Objekterkennung zu untersuchen.

Analyse der Reaktionszeiten der Gesamtstudie:

- deskriptive Analyse
- Verwendung von Maßen der zentralen Tendenz
  - lineares gemischtes Modell
- ohne Verwendung von Maßen der zentralen Tendenz
  - Vergleich der Verteilungen
  - Verteilungsdichte
  - empirische Verteilungsfunktion
- probabilistisches Modell
  - Ereigniszeitanalyse

## 11.1 Vergleich der Simulatorstudien

Zusammenfassend wurden in drei Studien die Leistungsfähigkeit der menschlichen visuellen Wahrnehmung untersucht. Vergleicht man die Ergebnisse der zwei Simulatorstudien miteinander, zeigen sich ähnliche Effekte in der Regressionsanalyse für die meisten variierten Einflussfaktoren. Unterschiede sind jedoch beim Faktor Zugsicherungssystem festzustellen. Dabei sind die Reaktionszeiten für FaS in der Simulatorstudie-BBI im Gegensatz zur DLR-Studie länger als die Bedingungen mit einem Zugsicherungssystem. Obwohl die Auswirkungen der unabhängigen Variablen in beiden Simulatoren ähnlich waren, sind die Reaktionszeiten in Simulatorstudie-BBI tendenziell länger als in Simulatorstudie-DLR. Dies könnte auf die Unterschiede in Bezug auf das Studiendesign und die Simulatorumgebung zurückzuführen sein. Erstens war der Versuchsaufbau unausgewogen, indem die BBI-Studie weniger Szenarien mit günstigen Bedingungen hatte (wie z. B. Bedingungen mit großen Objekten) (vgl. Abbildung 8).

Wenn nur die gemeinsamen Versuchsbedingungen des ETCS-Szenarios berücksichtigt werden, verringert sich der Unterschied zwischen den geometrischen Mitteln der BBI- und DLR-Reaktionszeiten von etwa 121 % auf 67 %. Die Differenz zwischen den geometrischen Mittelwerten der BBI- und DLR-Reaktionszeiten beträgt bei der PZB rund 78 %. Zweitens enthalten die Simulatoren Monitore mit unterschiedlicher Größe und Auflösung, die naturgemäß die Winkelgröße von Objekten beeinflussen, die auf den Bildschirmen dargestellt wurden. Die physikalische und wahrgenommene Größe des Reizes sowie die Bildschirmauflösung beeinflussen die Reaktionszeit bei Reizerkennungs- und visuellen Suchaufgaben (Plewan, 2021; Chen, Chiang 2012). Die Unterschiede könnten auf die etwas kleinere Darstellung der

Simulation auf einem Monitor (BBI) im Vergleich zur Darstellung mit einem Videoprojektor (DLR) zurückzuführen sein. Der Abstand zwischen Monitor und den Augen der Probanden ist ein weiterer wesentlicher Faktor, der die Wahrnehmung der Größe beeinflusst. Zur Berechnung der wahrgenommenen Größe der Stimuli in beiden Simulatoren wurde das Konzept der Winkelgröße verwendet. Die Winkelgröße repräsentiert den visuellen Winkel, den ein Objekt am Auge subtendiert, wobei sowohl die Objektgröße als auch der Betrachtungsabstand berücksichtigt werden. Es lässt sich folgende Beziehung zwischen der scheinbarer Größe  $\alpha$ , Entfernung  $D$  (Betrachtungsabstand) und den tatsächlichen Abmessungen eines Objekts  $g$ :

$$\alpha = 2 * \arctan\left(\frac{g}{2D}\right) \quad (G5)$$

Die Tabelle 18 zeigt die berechnete Winkelgröße der Stimuli für jeden Fahrsimulator. Die Objekte auf dem DLR-Simulator erscheinen im Vergleich zum Simulator-BBI etwa 70 % größer in Bezug auf die Winkelgröße. Dieser Unterschied liegt relativ nah an den oben berechneten Unterschieden der geometrischen Mittelwerte der Reaktionszeiten unter gemeinsamen Bedingungen. Dieser Unterschied in der Winkelgröße wurde bei der Analyse der Reaktionszeiten berücksichtigt.

TABELLE 18: SCHEINBARE GRÖÖE DER STIMULI AUF BBI- UND DLR-SIMULATOREN

Stimuli	ausgewählte Größe (cm)	scheinbare Größe (Winkelminute)
BBI_klein	90	3,78
BBI_groß	180	7,56
DLR_klein	90	6,72
DLR_groß	180	13,43

## 11.2 Deskriptive Analyse

Die Datenanalyse umfasste mehrere Schritte, um die erhobenen Daten zu untersuchen und aussagekräftige Ergebnisse zu erzielen. Die Zusammenführung der Datensätze von BBI und DLR ergab 702 Datenpunkte für die Gesamtanalyse. Zunächst wurde die Eigenschaft der kontinuierlichen Reaktionszeitdaten analysiert. Eine deskriptive Analyse wurde durchgeführt, um eine umfassende Zusammenfassung der Merkmale der Stichprobe zu erhalten.

Das Histogramm der Reaktionszeit zeigt erwartungsgemäß eine nicht-normalverteilte Datenverteilung mit einer positiven Schiefe (Abbildung 25, links). Wie bei den Analysen in Abschnitt 7.8 wurden die Rohreaktionszeiten in eine logarithmische Skala transformiert (Abbildung 25, links). Die kumulative Verteilungsfunktion (CDF) der transformierten Daten zeigt nahezu Normalität (Abbildung 25, rechts). Allerdings gibt es immer noch leichte Abweichungen von der roten Linie, die die angepasste Normalverteilung darstellt (insbesondere im Bereich der niedrigen Reaktionszeiten).

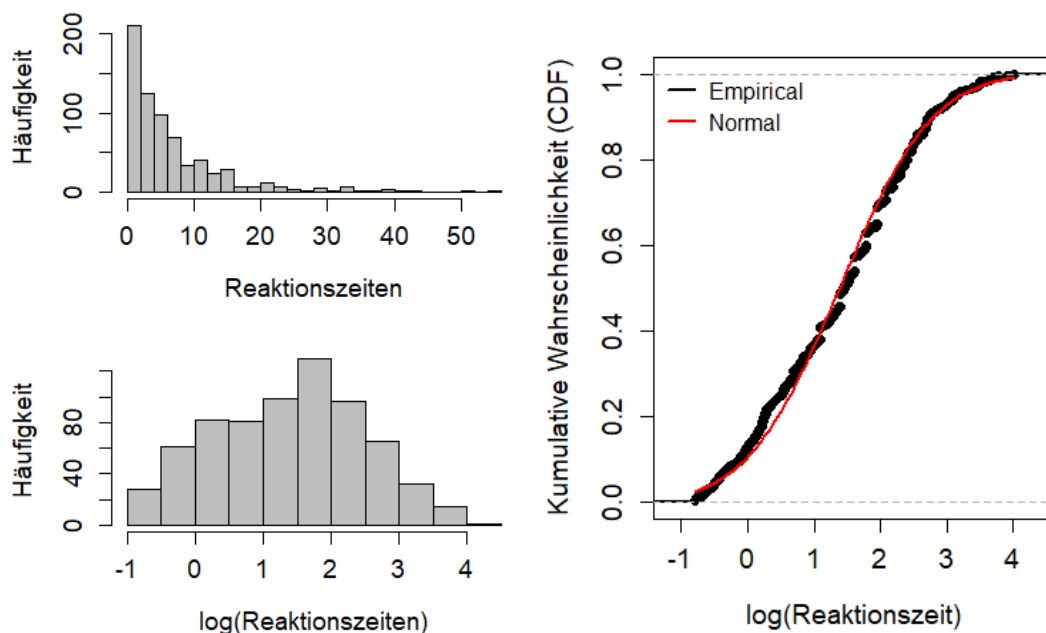


Abbildung 25: Histogramm der rohen und log-transformierten Rohreaktionszeiten (links) und empirische und theoretische kumulative Verteilungsfunktion der log-transformierten Reaktionszeitdaten (rechts)

Abbildung 26 zeigt die Maße der zentralen Tendenz auf dem Streudiagramm der Reaktionszeit für verschiedene Geschwindigkeitsstufen bei unterschiedlichen Zugsicherungssystemen. Die Zahl der statistischen Ausreißer (1,5x IQR) ist in der PZB am höchsten, die auch die größte Variabilität im Datensatz aufweist.

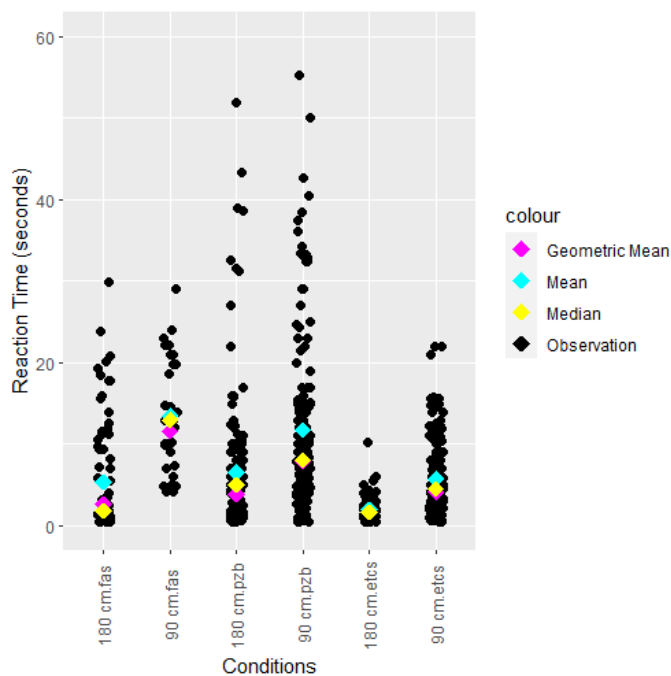


Abbildung 26: Maße der zentralen Tendenz einige Szenarien

Tabelle 19 zeigt eine Übersicht über die gemessenen geometrischen Mittelwerte und Standardabweichungen pro Stimulus. Da die Auswirkungen der Variablen in beiden Simulatorstudien größtenteils ähnlich waren, weisen die Ergebnisse der gemeinsamen Analyse auf die erwarteten Beziehungen hin. Höhere Geschwindigkeiten führen zu kürzeren Reaktionszeiten. Große Objekte und hoher Kontrast sind mit kürzeren Reaktionszeiten verbunden.

TABELLE 19: GEOMETRISCHE MITTELWERTE DER REAKTIONSZEIT DER VERSCHIEDENEN VARIABLENKOMBINATIONEN

	40 km/h			100 km/h			160 km/h		
<b>FaS</b>	kontrast-reich	groß:	2,42						
		klein:	9,92						
	kontrast-arm	groß:	2,78						
		klein:	13,44						
<b>PZB</b>	kontrast-reich	groß	4,11	kontrast-reich	groß	2,53			
		klein	11,84		klein	4,05			
	kontrast-arm	groß	7,21	kontrast-arm	groß	2,98			
		klein	13,69		klein	5,87			
<b>ETCS</b>				kontrast-reich	groß	1,45	kontrast-reich	groß	1,18
					klein	4,84		klein	3,04
				kontrast-arm	groß	2,13	kontrast-arm	groß	1,53
					klein	6,26		klein	3,16

Abbildung 27 zeigt die Reaktionszeiten für jede Objektgröße gemessen an der Winkelgröße (Bogenminute) anstelle der im Simulator gewählten Objektgröße (cm). Der Effekt der Größe ist erwartungsgemäß deutlich. Die Auswirkungen der einzelnen Variablen und ihr Einfluss auf die Reaktionszeit können mithilfe eines Regressionsmodells weiter analysiert werden.

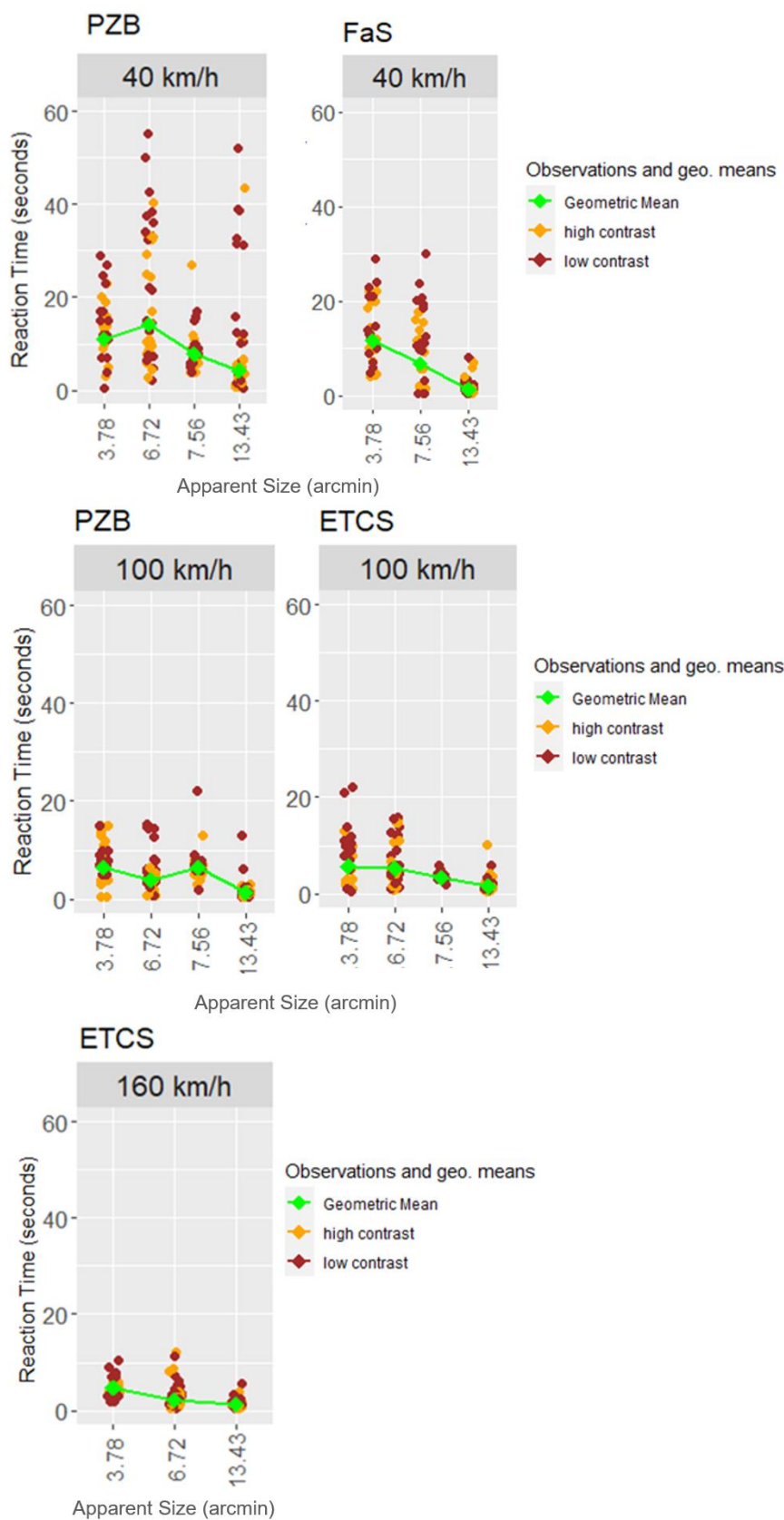


Abbildung 27: Reaktionszeiten für den gesamten Datensatz bei 40 km/h (oberes Feld), 100 km/h (mittleres Feld) und 160 km/h (unteres Feld). Y-Achse: Reaktionszeiten in Sekunden, X-Achse: Stimulusgröße in Bogenminuten. Die roten und gelben Punkte stellen die Beobachtungen für Stimuli mit niedrigem und hohem Kontrast dar.

## 11.3 Mixed-Effects-Modell

Für die Gesamtanalyse wurde ein Mixed-Effects-Modell verwendet, wobei der Simulator (BBI und DLR) als zufälliger Effekt berücksichtigt wurde. Durch die Berücksichtigung des zufälligen Effekts des Simulators können Unterschiede zwischen den Simulatoren erfasst werden, die möglicherweise einen systematischen Einfluss auf die Reaktionszeit haben. Dieser Ansatz ermöglicht es, sowohl die Gesamtauswirkungen der unabhängigen Variablen als auch die Variation zwischen den Simulatoren zu untersuchen.

### 11.3.1 Definition des Modells

Die festen Effekte stellen die Faktoren dar, die einen systematischen Einfluss auf die Reaktionszeit haben, während die zufälligen Effekte die Variabilität zwischen verschiedenen Simulatorumgebungen berücksichtigen. Obwohl eine größere Anzahl von Gruppen (größer als fünf) oft vorgeschlagen wird, um die Genauigkeit der Schätzung der Gruppenvarianz zu erhöhen, können Multilevel-Modelle selbst bei nur einer oder zwei Gruppen in den Daten nützlich sein, um Vorhersagen über neue Gruppen zu treffen (Gelman und Hill, 2009). Wenn jedoch die Niveaus eines Faktors innerhalb der Niveaus eines anderen Faktors variieren, wird dieser Faktor als verschachtelt bezeichnet. Da unterschiedliche Probanden an den Simulatorstudien teilnahmen, sind die Probanden in den Simulatoren verschachtelt. Für die lineare Regression wurde eine Mixed-Effects-Modellierung mit dem Ansatz der Restricted Maximum Likelihood (REML) verwendet. REML liefert unverzerrte Schätzwerte für die festen Effekte, indem die Likelihood-Funktion für die Komplexität des Modells bestraft wird (Searle, McCulloch und Casella, 1992).

Die funktionale Form des linearen Regressionsmodells mit Random-Intercepts (*M1*) lautet wie folgt:

$$\begin{aligned}
 & \text{Log(Reaktionszeit)} \\
 = & \beta_0 + \beta_1 \cdot \text{Größe} + \beta_2 \cdot \text{Kontrast} + \beta_{3,1} \cdot \text{Geschwindigkeit}_1 + \beta_{3,2} \cdot \text{Geschwindigkeit}_2 \\
 & + \beta_{4,1} \cdot \text{zsbs}_1 + \beta_{4,2} \cdot \text{zsbs}_2 + u_{\text{simulator}:i} + \varepsilon
 \end{aligned}
 \tag{G6}$$

$\beta_0$  repräsentiert den Achsenabschnitt oder den Basiswert der logarithmierten Reaktionszeit, wenn alle anderen unabhängigen Variablen auf ihren Referenzwerten gehalten wurden. Zusätzlich zu den Parametern für die festen Effekte  $\beta_i$  für die Kovariaten, die bei der Konstruktion der Designmatrix verwendet wurden, enthält das Modell zwei Zufallskomponenten: die gruppeninternen Restfehler  $\varepsilon_i$  und die Zufallseffekte  $u_j$  für die Kovariaten. Die Restfehler für dieselbe Gruppe sind unabhängig von den Zufallseffekten. Der Zufallsfaktor berücksichtigt die Unterschiede in den Abschnitten zwischen den beiden Simulatoren, die auf inhärente Unterschiede oder unbeobachtete Faktoren in Bezug auf jeden Simulator zurückzuführen sind. In dem zufälligen Teil des Modells sind die Parameter, die wir schätzen, die Varianzen  $\sigma_u^2$  und  $\sigma_\varepsilon^2$ .

Tabelle 20 enthält die Dummy-Variablen, die als Indikator für das Vorhandensein einer Ausprägung einer mehrstufigen Variablen dienen. Das Referenzniveau ist: PZB, 100 km/h, großer Stimulus, kontrastreicher Stimulus. Es ist wichtig zu beachten, dass die Wahl des Referenzniveaus keinen Einfluss auf die allgemeine Anpassung des Modells oder die Gültigkeit der statistischen Analyse hat. Sie beeinflusst hauptsächlich die Interpretation der Modellkoeffizienten und die Vergleiche zwischen den Ebenen.

TABELLE 20: KODIERUNG VON DUMMY-VARIABLEN FÜR ALLE FAKTOREN

ZSS	zss1	zss2	Geschwindigkeit	geschw.1	geschw.2	Größe	größe	Kontrast	kontrast
PZB	0	0	100 km/h	0	0	groß	0	kontrastreich	0
FaS	1	0	40 km/h	1	0	klein	1	kontrastarm	1
ETCS	0	1	160 km/h	0	1				

Die in den Kapiteln 8 und 9 definierten Forschungshypothesen für Zweistufenvariablen gelten hier ebenfalls. Bei Dreistufenvariablen wird die Referenzvariable mit den anderen beiden Variablen verglichen. Bei zss1 wird erwartet, dass die Reaktionszeiten bei FaS niedriger sind als bei PZB, und für die erste Geschwindigkeitsvariable (*geschwindigkeit1*) wird erwartet, dass die Reaktionszeit bei 100 km/h niedriger ist als bei 40 km/h.

### 11.3.2 Ergebnisse des Modells mit gemischten Effekten

Diese Studie untersucht die Faktoren, die die Reaktionszeit durch ein Mixed-Effects-Lineares-Regression-Modell beeinflussen. Das R-Paket *lme4* (Bates et al., 2015) wurde zur Durchführung einer Analyse linearer Mixed-Effekte verwendet. Das Modell wurde auf einen Datensatz mit 702 Beobachtungen angewendet, wobei die abhängige Variable die logarithmierte Reaktionszeit war. Die festen und zufälligen Effekte wurden analysiert, um ihren Einfluss auf die Reaktionszeit zu bewerten. Der Konditional -R<sup>2</sup> betrug 0.52, was darauf hindeutet, dass das gesamte Modell etwa 52 % der Varianz in der Reaktionszeit erklärt.

Die Schätzungen und Modellstatistiken wurden in der Tabelle 21 präsentiert. Jeder Koeffizient für den fixen Effekt repräsentiert die erwartete Veränderung in der logarithmierten Reaktionszeit bei einer Einheit Veränderung in der entsprechenden Prädiktorvariable. Der Standardfehler (SF) liefert eine Schätzung der Unsicherheit des Koeffizienten. Der t-Wert und der p-Wert bewerten die statistische Signifikanz des Koeffizienten.

TABELLE 21: SCHÄTZUNGEN DES MODELLS MIT GEMISCHTEN EFFEKTEN (M1) UND MODELLPARAMETER

Feste Effekte				Zufällige Effekte					
	Koeffizient	Exp (ß)	SF	t	5 %	95 %	Group	Varianz	SD
Intercept*	0,79	2,20	0,11	7,49	0,58	0,99	Simulator (intercept)	0,26	0,49
größe_90 (ß1)*	0,82	2,27	0,06	13,92	0,70	0,93	Residual	0,55	0,74
kontrast_niedrig (ß2)*	0,22	1,25	0,06	3,91	0,11	0,33	Random-Intercept	BBI	DLR
geschw._40 (ß3.1)*	0,83	2,29	0,08	9,96	0,67	1,00		1,14	0,44
geschw._160 (ß3.2)*	-0,39	0,68	0,09	-4,29	-0,56	-0,21	Modellparameter		
fas (ß4.1)*	-0,68	0,51	0,09	-7,23	-0,86	-0,49	AIC	1702	
etcs (ß4.2) * p<0.05	-0,10	0,90	0,09	-1,18	-0,27	0,07	BIC	1748	
							ICC	0,31	
							Pseudo-R <sup>2</sup> (total)	0,58	

P-Werte für gemischte Modelle sind nicht so intuitiv zu interpretieren wie für das lineare Modell. Die genauen Nullverteilungen für die Parameterschätzungen und Teststatistiken sind unbekannt (Bates et al., 2015). Daher ist die einzige Möglichkeit, die Signifikanz der Effekte zu beurteilen, eine Methode der Annäherung. In dieser Analyse wurden die einseitigen p-Werte unter Verwendung von Satterthwaite-Methode berechnet. Der p-Wert der Variable *etcs* legt nahe, dass diese Beziehung zwischen PZB und ETCS statistisch nicht signifikant ist. Schaalje, McBride und Fellingham (2002) haben jedoch gezeigt, dass die Komplexität der Kovarianzstrukturen, der Stichprobenumfang und das Ungleichgewicht die Leistung der Approximationen beeinflussen.

$$\text{Log(Reaktionszeit)} = 0,79 + 0,82 \cdot \text{größe}_{\text{klein}} + 0,22 \cdot \text{kontrast}_{\text{kontrastarm}} + 0,83 \cdot \text{geschwindigkeit}_{40\text{km/h}} - 0,39 \cdot \text{geschwindigkeit}_{160\text{km/h}} - 0,68 \cdot \text{zbs}_{\text{fas}} - 0,10 \cdot \text{zbs}_{\text{etcs}} + u_{\text{simulator}} \quad (\text{G7})$$

Die Interpretation der Koeffizienten ist identisch mit der der normalen Regressionsmodelle. Die positiven und negativen Vorzeichen zeigen die Erhöhung oder Verringerung der Reaktionszeit im Vergleich zum Basisniveau an. Die potenzierten Koeffizienten können interpretiert werden, um die multiplikative Auswirkung einer Einheit Veränderung in der entsprechenden Prädiktorvariable auf die Reaktionszeit zu verstehen. Zum Beispiel ist die erwartete Reaktionszeit bei einer Geschwindigkeit von 40 km/h um 129 % größer als bei einer Geschwindigkeit von 100 km/h. Andererseits ist die erwartete Reaktionszeit bei einer Geschwindigkeit von 160 km/h um 59 % geringer als bei einer Geschwindigkeit von 100 km/h. Alle Hypothesen außer der Beziehung zwischen ETCS und PZB können angenommen werden.

Die zufälligen Effekte in diesem Modell tragen zur Berücksichtigung der Variabilität in den Daten aufgrund der Gruppierungsstruktur bei. Die Unterschiede zwischen den geometrischen Mittelwerten der Reaktionszeiten zwischen zwei Simulatoren betragen 67 %, wenn nur die gleichen Versuchsbedingungen berücksichtigt wurden. Das zufällige Intercept der Simulatoren hat eine Varianz von 0.26, was auf eine erhebliche Variabilität in der Reaktionszeit zwischen den verschiedenen Simulatoren hinweist. Die geschätzte Standardabweichung dieses Zufallseffekts beträgt, gemessen in log-transformierter Skala, etwa 0,49. Der geschätzte klasseninterne Korrelationskoeffizient (ICC) für diese Gruppierungsvariable beträgt 0,31. Etwa 31 % der Gesamtvariation in der Reaktionszeit waren also auf Unterschiede zwischen den Simulatorgruppen zurückzuführen. Die Interpretation des Achsenabschnitts weicht leicht von der normalen linearen Regression ab. Es kann ein Achsenwert pro Probanden oder pro Simulatorgruppe berechnet werden, um die Unterschiede zwischen den Simulatoren zu berücksichtigen.

Es wird angenommen, dass die zufälligen Effekte normalverteilt sind und der Mittelwert Null beträgt. Oft ist es sinnvoll, die Probanden als zufällig aus einer größeren Population ausgewählt zu betrachten, deren Merkmale abgeschätzt werden sollen. Daher ist es sinnvoller, die Faktoren zu ermitteln, die die Verteilung der Zufallseffekte beschreiben, als die Schätzungen der Zufallseffekte. Die Abbildung 28 zeigt die Verteilung der Achsenabschnitte für Probanden, die in Simulatoren geclustert wurden. Diese Clusterung deutet darauf hin, dass die Intercepts der BBI-Teilnehmenden im Vergleich zu den DLR-Teilnehmenden zu höheren Werten verschoben sind. Wie im Kapitel 11.1 dargestellt, war die geometrische mittlere Reaktionszeit in der BBI-Studie tendenziell länger als der DLR-Studie (78 % bei PZB, 67 % bei ETCS). Ein möglicher Grund dafür könnte die Größe und Auflösung der Monitore sein (Tabelle 18). Um die Auswirkung der physischen Größe auf die Varianz des Achsenabschnitts zu bewerten, wurde ein weiteres Modell (*M2*) an die Daten angepasst, indem die Objektgröße als Winkelgröße definiert wurde, die für jede Simulatorstellung spezifisch ist.

$$\text{Log(Reaktionszeit)} = \beta_0 + \beta_{1.1} \cdot \text{Winkelgröße}_1 + \beta_{1.2} \cdot \text{Winkelgröße}_2 + \beta_{1.1} \cdot \text{Winkelgröße}_3 + \beta_2 \cdot \text{Kontrast} + \beta_{3.1} \cdot \text{Geschwindigkeit}_1 + \beta_{3.2} \cdot \text{Geschwindigkeit}_2 + \beta_{4.1} \cdot \text{zss}_1 + \beta_{4.2} \cdot \text{zss}_2 + u_{\text{simulator:teilnehmenden}} + \varepsilon \quad (\text{G8})$$

In der Modellnotation M2 wurden drei Dummy-Variablen für die Winkelgröße der Objekte definiert (Tabelle 22). Der erste Begriff steht für den Vergleich der Gesamtunterschiede zwischen DLR- und BBI-Bedingungen. Der zweite und dritte Begriff bezeichnen die paarweisen Vergleiche von kleinen und großen Objekten für BBI- bzw. DLR-Simulatoren. Indem man die Unterschiede in der Größe der BBI- und DLR-Objekte als feste Effekte einbezieht, kann man den systematischen Effekt dieses Unterschieds aus der Zufallsvarianz extrahieren.

TABELLE 22: DUMMY-VARIABLE FÜR DIE WINKELGRÖÖE UND DIE PAARWEISEN VERGLEICHE DURCH KONTRASTKODIERUNG

Dummy-Variable	Vergleich
<i>Winkelgröße</i> <sub>1</sub>	BBI - DLR
<i>Winkelgröße</i> <sub>2</sub>	groß - klein (BBI)
<i>Winkelgröße</i> <sub>3</sub>	groß - klein (DLR)

Die Ergebnisse der Analyse des linearen gemischten Effekts sind in der Tabelle 23 aufgeführt. Der potenzierte Koeffizient für *Winkelgröße*<sub>1</sub> deutet darauf hin, dass der Unterschied zwischen den Simulatoren 43 % beträgt, wenn alle anderen Bedingungen auf ihrem Referenzniveau liegen.

TABELLE 23: SCHÄTZUNGEN DES MODELLS MIT GEMISCHTEN EFFEKTEN (M2)

Feste Effekte						
	Koeffizient	Exp (β)	SF	t	2.5%	97.5%
<b>Intercept*</b>	1.26	3.52	0.09	14.51	1.09	1.44
<b>Winkelgröße1(β1.1)*</b>	0.36	1.43	0.07	5.50	0.49	0.22
<b>Winkelgröße2(β1.2)*</b>	0.18	1.20	0.06	3.91	0.26	0.09
<b>Winkelgröße3 (β1.3)*</b>	0.54	1.72	0.05	14.96	0.61	0.48
<b>Kontrast_niedrig (β2)*</b>	0.20	1.22	0.05	3.64	0.09	0.31
<b>Geschwindigkeit_40 (β3.1)*</b>	0.85	2.34	0.08	10.37	0.69	1.01
<b>Geschwindigkeit_160 (β3.2)*</b>	-0.34	0.71	0.09	-3.91	-0.52	-0.16
<b>fas (β4.1)*</b>	-0.63	0.53	0.09	-6.88	-0.80	-0.45
<b>etcs (β4.2)</b>	-0.06	0.94	0.08	-0.72	-0.22	0.09

\* p<0.05

Der Vergleich der Modellstatistiken ist in der Tabelle 24 dargestellt. Es ist zu erkennen, dass die geschätzte Varianz des zufälligen Achsenabschnitts von 0,26 auf 0,14 verringert wurde. Der Interklassen-Korrelation-Wert (ICC) zeigt, dass die durch die Clusterunterschiede erklärte Varianz von 32 % auf 20 % reduziert wurde. Dies bedeutet, dass ein Großteil der Varianz zwischen den Zufallsabschnitten auf die physikalische Objektgröße zurückzuführen sein könnte. Die Abbildung 28 (rechts) zeigt die zufälligen Achsenabschnitte der einzelnen Probanden innerhalb jedes Simulators. Die Varianz, die durch die Unterschiede in der Winkelgröße verursacht wurde, wurde durch die festen Effekte erfasst. Im Vergleich zu den zufälligen Achsenabschnitten von Modell-1 wurden die Clustering-Effekte deutlich abgeschwächt (vgl. Abbildung 28, links). R-Quadrat-Werte zeigen eine Verbesserung der Modellanpassung in Bezug

auf systematische Effekte an. Ein niedrigeres Bayes'sches Informationskriterium (BIC), das die Modellkomplexität bestraft, deutet ebenfalls auf eine bessere Anpassung hin.

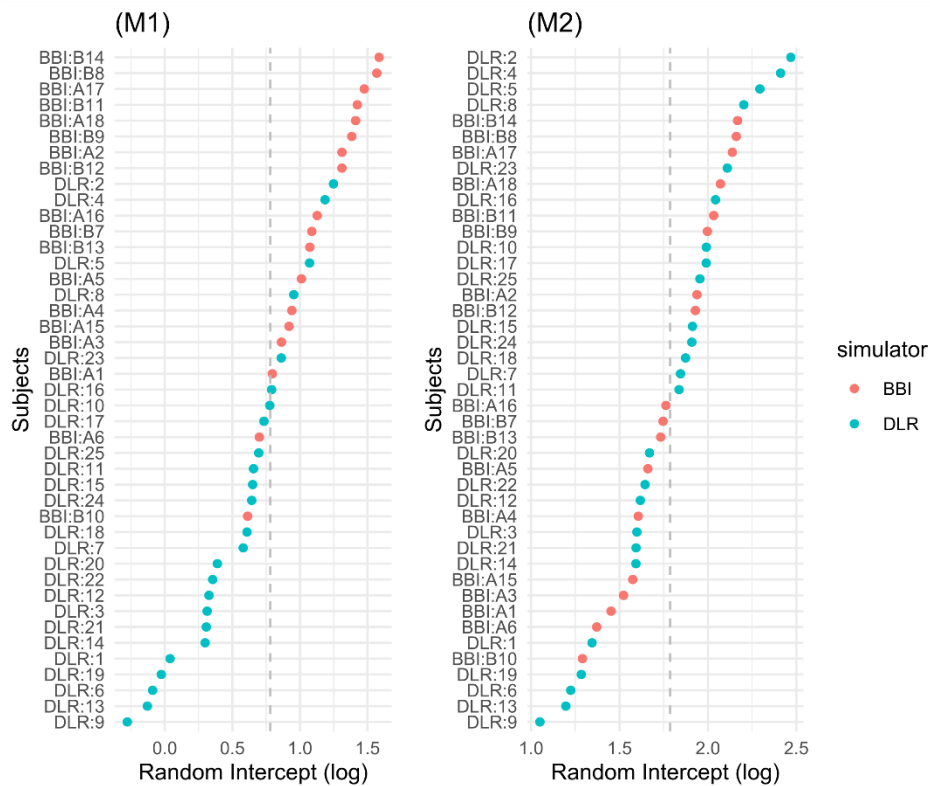


Abbildung 28: Einzelne Achsenabschnitte von Probanden, die innerhalb jedes Simulators geclustert wurden (Modell-1) (links), Einzelne Achsenabschnitte, wenn die Unterschiede in der Winkelgröße berücksichtigt wurden. (Modell-2) (rechts).

Die individuellen Achsenabschnitte der Probanden sind in der Abbildung 29 dargestellt.

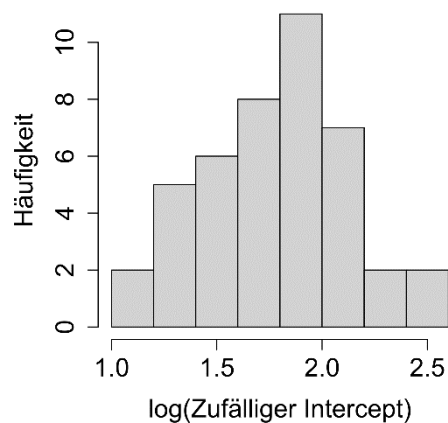


Abbildung 29: Histogramm der einzelnen Achsenabschnitte der Probanden (M2).

TABELLE 24: VERGLEICH DER ZUFALLSEFFEKTE UND MODELLPARAMETER DER MODELLE M1 UND M2

	Modell-1	Modell-2
<b>Varianz (Simulator)</b>	0.26	0.14
<b>Varianz (Residuen)</b>	0.55	0.53
<b>ICC-Simulator</b>	32 %	20 %
<b>SD-Simulator: Teilnehmenden</b>	0.49	0.37
<b>Pseudo-R<sup>2</sup> (fixed)</b>	0.30	0.45
<b>Pseudo-R<sup>2</sup> (total)</b>	0.58	0.57
<b>AIC-BIC</b>	1702 - 1748	1654 - 1704

### 11.3.3 Modellbewertung und -validierung

Die Modelldiagnose wurde durchgeführt, um die Annahmen zu bewerten und einflussreiche Daten für das Modell mit gemischten Effekten zu ermitteln. Residuen sind die Differenz zwischen den erhobenen und den angepassten Werten. Bei gemischten Effekten können die Residuen als Schätzungen von  $\epsilon$  betrachtet werden (Faraway, 2016). Es wurde angenommen, dass die Fehlervarianz konstant ist. Die Darstellung der Residuen im Vergleich zu den angepassten Werten zeigte eine zufällige Streuung der Residuen um Null, was auf keine besonderen Probleme hinsichtlich der Linearitätsannahme hinweist ( , rechts). Asymmetrisch verteilte Residuen am unteren Ende der angepassten Werte können jedoch auf Ausreißer oder einflussreiche Punkte in den Daten hinweisen. Es gibt auch keine deutlichen Hinweise auf Multikollinearität zwischen den unabhängigen Variablen im Modell (maximaler VIF = 2,88, durchschnittlicher VIF = 1,94). Der QQ-Plot zeigt den Vergleich der beobachteten Quantile mit den erwarteten Quantilen unter der Annahme einer Normalverteilung. Die Mehrheit der Punkte im QQ-Plot folgt der 45-Grad-Linie, was darauf hinweist, dass die Residuen des Modells annähernd normalverteilt sind ( , links). Am unteren Ende der Linie gibt es einige Abweichungen nach unten, was darauf hindeutet, dass der linke Randbereich der Verteilung länger ist, als unter Normalität erwartet. Insgesamt konnten die Ergebnisse der Modelldiagnose die Modellannahmen erfüllen, könnten jedoch auf Modellanpassungsprobleme bei Extremwerten hinweisen.

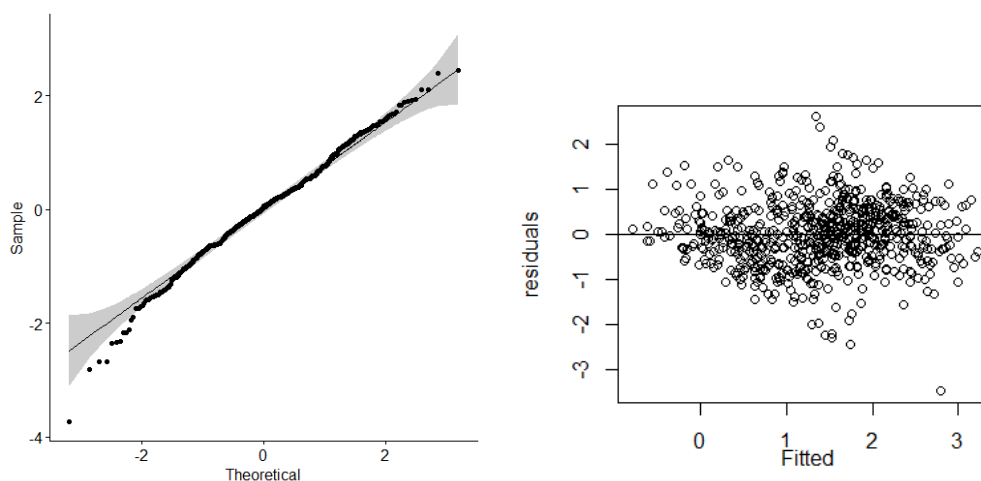


Abbildung 30: QQ-Plot zur Darstellung des Vergleichs der beobachteten Quantile mit den erwarteten Quantilen (links), Diagramm der Residuen und der angepassten Werte (rechts)

Außerdem wurde der Cook's-Abstand berechnet, um einflussreiche Beobachtungen zu identifizieren. Die Cook-Methode misst die Veränderung der Modellanpassung durch den Ausschluss jeder einzelnen Beobachtung, wobei sowohl der Abstand (Veränderung der vorhergesagten Werte) als auch die Hebelwerte (Einfluss der Beobachtung auf die Parameterschätzungen) berücksichtigt werden (Cook, 1977). Beobachtungen mit hohen Cook-Distanz-Werten zeigen an, dass ihre Entfernung zu einer wesentlichen Veränderung der Regressionskoeffizienten führen würde und somit ihren Einfluss auf das gesamte Modell verdeutlicht. In der Literatur finden sich verschiedene Schwellenwerte für den Cook-Abstand wie 0,5, 1 oder  $4/n$  (wobei  $n$  die Anzahl der Daten ist) (Cook, 2011). Bei einem Schwellenwert von 0,5 wurden keine Ausreißer festgestellt. Es wurde festgestellt, dass es 35 Ausreißer gab, bei denen der Cook-Abstand den Schwellenwert von  $4/n$  überschritt, und zwar von insgesamt 702 Datenpunkten (Abbildung 31). Um Ausreißer aus dem Datensatz zu eliminieren, müssen sie eindeutig identifiziert werden, was sehr kompliziert ist. Ratcliff (1993) empfiehlt, die Auswirkungen potenzieller Ausreißer zu reduzieren und dabei so wenig wie möglich von den eigentlich relevanten Daten zu eliminieren. Die Eliminierung von Ausreißern mit solchen Cut-off-Methoden kann jedoch die tatsächlichen Auswirkungen verschleiern (Sanders, 1998). In dieser Studie wurden die Ausreißer bei der Reaktionszeit in die Analyse einbezogen und ihr potenzieller Einfluss wird diskutiert.

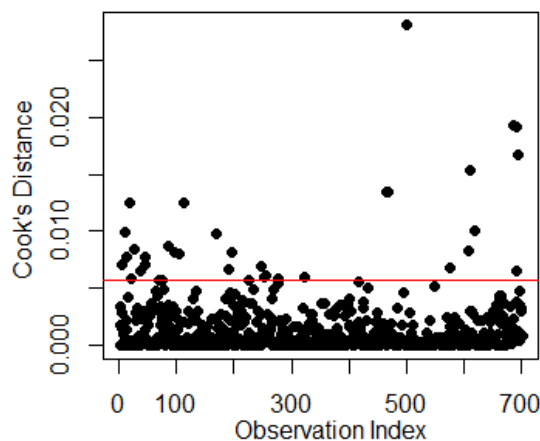


Abbildung 31: Cook's-Distanz-Plot zur Darstellung des Einflusses einzelner Beobachtungen auf das Regressionsmodell. Die rote Linie stellt den Schwellenwert ( $4/n$ ) dar.

Um die Wirksamkeit von Zufallseffekten zu bewerten, kann der Likelihood-Ratio-Test (LRT) mit der parametrischen Bootstrap-Methode verwendet werden (Faraway, 2016). Die LRT-Methode testet die Nullhypothese, dass die Varianz zwischen den Modelltermen null ist. Im Modell mit gemischten Effekten wurde das Subjekt, das innerhalb des Simulators geclustert wurde, als Zufallsbegriff aufgenommen, so dass die Hypothese, dass diese Varianz null ist, behauptet, dass es keinen Unterschied zwischen den Modellen gibt. Die LRT auf Basis von 10.000 simulierten Werten deutet darauf hin, dass der Begriff des Zufallseffekts erheblich zur Anpassung des Modells beiträgt (RLRT = 97,7,  $p < 0,01$ ).

Wenn man die Varianz über zufällige Effekte aus dem Modell herausnimmt, würden die festen Effekte zuverlässiger. Dies wird durch den Vergleich der Standardfehler der einzelnen Variablen bei der linearen Regression mit und ohne Zufallseffekte untersucht (Tabelle 25). Die Standardfehler der Mehrheit der Schätzungen mit festen Effekten wurden durch die Einbeziehung von Zufallseffekten geringfügig verbessert. Außerdem ist der AIC-Wert (1661) des Modells mit Zufallseffekt kleiner als der des Modells ohne Zufallseffekt (1843). Dies deutet auf eine bessere Anpassung des Modells mit zufälligen Effekten hin. Außerdem wurden die unteren und oberen Perzentile der Bootstrap-Verteilung mithilfe der Efron-Perzentilmethode geschätzt, um konservativere Konfidenzintervalle zu konstruieren. In der Tabelle 25

sind die mit Bootstrap berechneten Intervalle für das Modell aufgeführt. Die Bootstrap-Stichproben wurden durch wiederholtes Ziehen von 1.000 Simulationen aus den Daten generiert, und für jeden neu gezogenen Datensatz wurden die geschätzten Koeffizienten berechnet.

TABELLE 25: STANDARD-FEHLER (MIT UND OHNE RANDOM-EFFEKTE)

	SF_mit	SF_ohne
<b>size90 (<math>\beta_1</math>)</b>	0.06	0.07
<b>contrast-low (<math>\beta_2</math>)</b>	0.05	0.07
<b>speed40 (<math>\beta_{3.1}</math>)</b>	0.08	0.10
<b>speed160 (<math>\beta_{3.2}</math>)</b>	0.09	0.11
<b>fas (<math>\beta_{4.1}</math>)</b>	0.09	0.11
<b>etcs (<math>\beta_{4.2}</math>)</b>	0.09	0.10

Um das Modell zu validieren, wurden die Daten nach dem Zufallsprinzip in Trainings- und Validierungssätze aufgeteilt. Die Modellmetriken mittlerer quadratischer Fehler und R-Quadrat wurden sowohl für die Trainings- als auch für die Validierungssätze berechnet. Das Modell erfasst einen moderaten Anteil der Varianz in den Daten, aber es gibt immer noch einen beträchtlichen Teil der Varianz, der unerklärt bleibt. Die Ähnlichkeit der MSE- und R-Quadrat-Werte deutet darauf hin, dass das Modell die Trainingsdaten nicht signifikant übererfüllt und in der Lage ist, relativ gut auf die Validierungsdaten zu generalisieren (Tabelle 26). Es ist jedoch nicht möglich, allein auf der Grundlage der bereitgestellten Metriken endgültig festzustellen, ob das Modell ein Overfitting aufweist. Um eine Überanpassung festzustellen, sollte die Leistung des Modells anhand eines unabhängigen Testdatensatzes bewertet werden, der nicht für das Training oder die Validierung des Modells verwendet wurde.

TABELLE 26: MODELLPARAMETER DER TRAININGS- UND VALIDIERUNGSDATENSÄTZE

	Training	Validation
<b>Mittlerer Quadratfehler (MSE)</b>	0,498	0,518
<b>R<sup>2</sup></b>	0,587	0,551

### 11.3.4 Vorhersage

In diesem Abschnitt wurde der Prozess der Vorhersage im Mixed-Effects-Modell erläutert. Es ist jedoch wichtig zu beachten, dass gemischte Effektmodelle sowohl feste als auch zufällige Effekte berücksichtigen. Zufallsfaktoren tragen auf verschiedene Weise zur Vorhersage bei. Marginale Vorhersagen schätzen die durchschnittliche Reaktionszeit über alle Ebenen der zufälligen Effekte hinweg, während konditionale Vorhersagen die spezifischen zufälligen Effekte berücksichtigen, die mit jedem einzelnen Fall verbunden sind (Welham et al., 2004).

Marginale Vorhersagen bieten eine Gesamtschätzung der abhängigen Variablen, wobei die Variabilität durch die zufälligen Effekte ignoriert wird. Sie sind nützlich, wenn der Schwerpunkt auf den festen Effekten liegt und die durchschnittliche Antwort über alle Ebenen der zufälligen Effekte von Interesse ist. Bei dieser Methode wird zwar die durch Zufallseffekte erklärte Varianz nicht in die Vorhersage einbezogen, aber die Schätzungen der festen Effekte haben die Zufallseffekte bei der Erstellung unverzerrter

Schätzungen bereits berücksichtigt. Konditionale Vorhersagen bieten gruppenspezifische Schätzungen, indem sie die zufälligen Effekte berücksichtigen. Die zufälligen Abschnitte ermöglichen es dem Modell, die Abweichungen von der durchschnittlichen Beziehung für jede Gruppe zu erfassen. Diese Vorhersagen berücksichtigen den spezifischen Level der zufälligen Effekte, der mit jedem Fall verbunden ist, und erfassen somit die individuelle Variabilität.

Marginalmodelle wurden oft verwendet, wenn das Ziel darin besteht, populationsbasierte Aussagen zu machen (Diggle et al., 2013). Lee und Nelder (2004) widersprechen, indem sie sagen, dass der marginale Mittelwert nur dann mit dem Populationsdurchschnitt identisch sein kann, wenn die verwendeten Cluster repräsentative Zufallsstichproben aus der Population sind. Wenn man beispielsweise annimmt, dass die an der Studie teilnehmenden Probanden repräsentative Zufallsstichproben der realen Welt sind, können wir für die Vorhersagen das Marginalmodell verwenden. Eine Vorhersage des Populationsmittels (bzw. Marginal) kann verwendet werden, um eine Schätzung für diejenigen zu liefern, die nicht zu einem der bei der Modellentwicklung verwendeten Cluster gehören (Pavlou et al., 2015). In diesem Fall hat jedoch eine Beobachtung aus einem neuen Cluster (z. B. Simulator oder reale Welt) einen unbekanntem Zufallseffekt, aber je nach Ziel der Studie könnte nur der systematische (bzw. feste) Effekt wichtiger sein.

Die zufälligen Intercepts der Probanden sind in Abbildung 29 und Tabelle 24 dargestellt. In dieser Studie wurden sowohl die marginale Vorhersage (Abbildung 33) als auch die konditionale Vorhersage für jeden Probanden (Abbildung 32) berechnet.

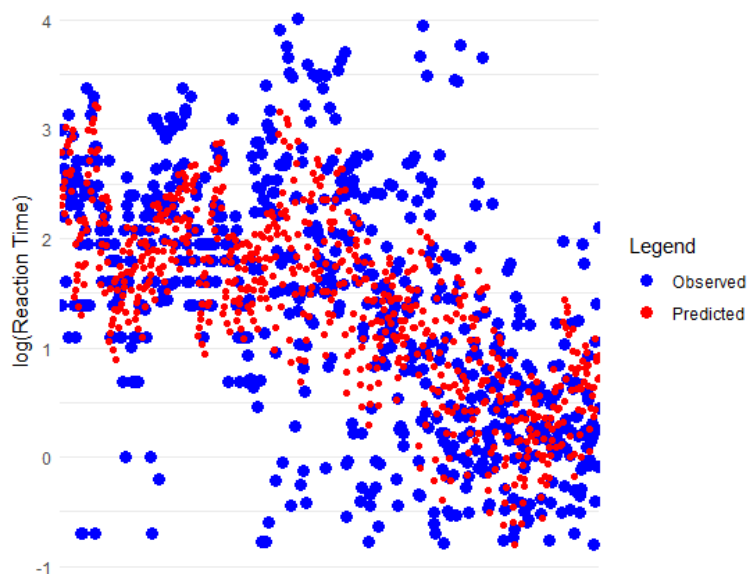


Abbildung 32: Beobachtete Datenpunkte und die Modellvorhersagen basierend auf den individuellen Achsenabschnitten der Probanden für jeden Stimulus. X-Achse: individuelle Teilnehmenden-ID.

Das geometrische Mittel der beobachteten Daten und der Vorhersagen ist in der Abbildung 33 dargestellt. Die Schätzungen wurden für vier verschiedene Größenklassen abgegeben, die als Winkelgröße des Objekts in jedem Simulator gemessen wurden.

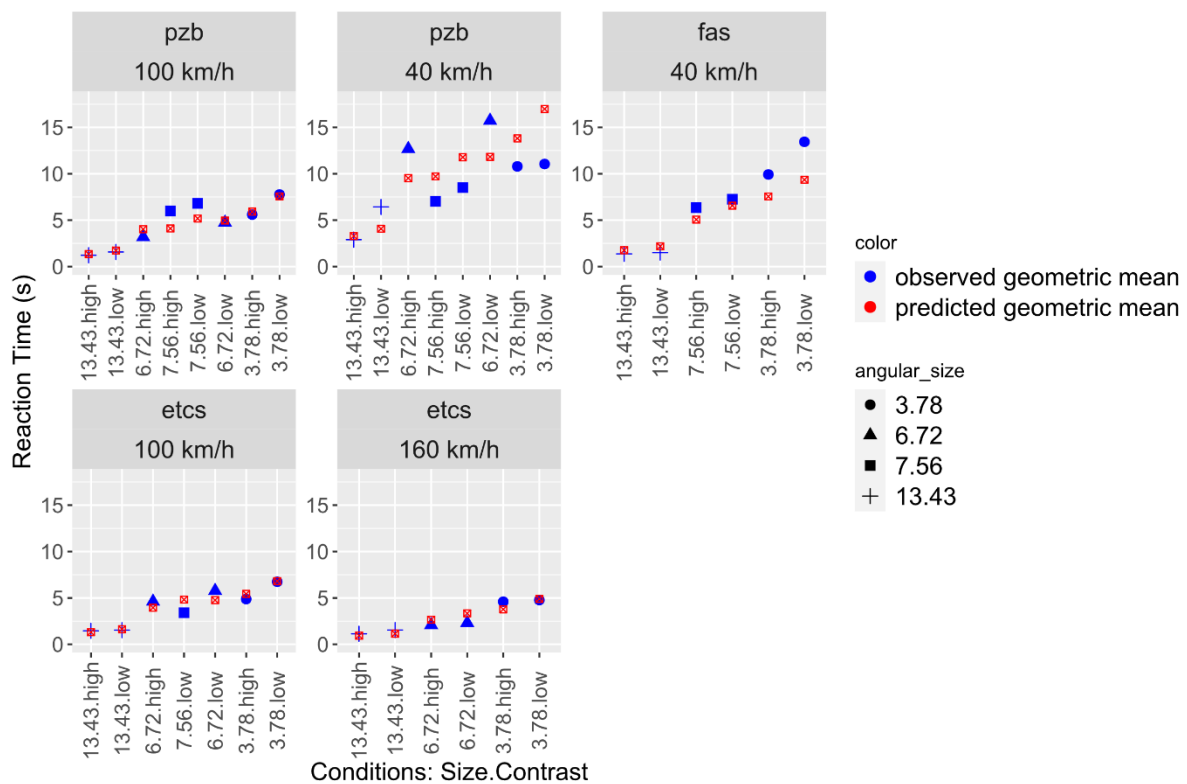


Abbildung 33: Vergleich der Modellvorhersage mit dem entsprechenden Indikator der zentralen Tendenz über die Bedingungen hinweg. Y-Achse: Reaktionszeiten in Sekunden. X-Achse: Versuchsbedingungen als Kombination von Winkelgröße und Kontraststufen.

Die Wahl des verwendeten Modells hängt jedoch von den Annahmen und dem Ziel der Vorhersagen ab. Es ist zu beachten, dass bei den Experimenten in dieser Studie spezifische Szenarien und Randbedingungen verwendet wurden. Jede Vorhersage, die auf der Grundlage dieses Modells getroffen wird, muss diese Besonderheiten berücksichtigen. Außerdem sollte der Anteil der Varianz, der durch das Modell erklärt wird, berücksichtigt werden. Eine Empfehlung zur weiteren Verbesserung des Modells wäre, einen neuen Datensatz von den Mitgliedern eines neuen Clusters zu erhalten, um das Modell damit zu kalibrieren.

### 11.3.5 Auswertungstool

Im Rahmen des AP4 wurde das Modell in ein Tool zur Ableitung von Anforderungen überführt. Dieses Tool wurde mithilfe des R-Tool entwickelt und ermöglicht eine automatisierte Analyse sowie die anschauliche Darstellung der Ergebnisse. Es bietet die Möglichkeit, die Ergebnisse des Regressionsmodells sowie vorhergesagte Reaktionszeiten und Wahrscheinlichkeiten unter verschiedenen, vom Benutzer festgelegten Bedingungen zu visualisieren. Des Weiteren ermöglicht das Tool durch Interpolation die Vorhersage von Reaktionszeiten und Wahrscheinlichkeiten für Bedingungen, die nicht explizit getestet wurden. Die Genauigkeit der Vorhersagen kann durch die Vereinfachungen im Modell sowie durch die Annahmen, die mit der Regression verbunden sind, beeinflusst werden. Daher sollte das Tool vorrangig für Demonstrationszwecke verwendet werden und nicht als allgemeine Grundlage für technische Anforderungen dienen.

Unter Verwendung einer angepassten Version des Modell-2 (G8) können Schätzungen der Reaktionszeit für verschiedene Fahrgeschwindigkeiten und für verschiedene Werte der Winkelgröße des Objekts ausgegeben werden. Die tatsächlichen Geschwindigkeitswerte, die im Moment des Auftretens des Objekts gemessen wurden, wurden im Datensatz eingetragen. Zusätzlich wurde die Variable „Winkelgröße“ als kontinuierliche Variable in das Modell aufgenommen. Objektkontrast zum Hintergrund und Zugsicherungssysteme wurden als diskrete Variablen mit den gleichen Ebenen wie im Modell-2 (G8) aufgenommen. Die funktionale Form des linearen Regressionsmodells (M3) lautet wie folgt:

$$\text{Log(Reaktionszeit)} = \beta_0 + \beta_1 \cdot \text{Winkelgröße} + \beta_2 \cdot \text{Kontrast}_{\text{hoch}} + \beta_3 \cdot \text{Geschwindigkeit} + \beta_{4.1} \cdot \text{zss}_{\text{fas}} + \beta_{4.2} \cdot \text{zss}_{\text{etcs}} + u_{\text{simulator:teilnehmenden}} + \varepsilon \quad (\text{G9})$$

Das Auswertungstool ermöglicht die Eingabe von Benutzerwerten für die kontinuierlichen unabhängigen Variablen und Ebenen für kategoriale Variablen. Die Objektgröße wird in Zentimetern eingegeben, das Tool wandelt sie in einen Sehwinkel in Bogenminuten um, unter der Annahme einer Entfernung des Beobachters zum Objekt von 800 Metern. Der Mindestwert für die Eingabe beträgt 30 cm, was etwa dem minimalen Sehwinkel des menschlichen Auges mit normaler Sehschärfe in einer Entfernung von 800 Metern entspricht. Der Eingabebereich für die Fahrgeschwindigkeit liegt zwischen 15 km/h und 180 km/h. Der maximale Wert für die Fahrgeschwindigkeit in einem FAS-Szenario beträgt 40 km/h. Unter Verwendung einer multiplen linearen Regression wurden die Koeffizienten jeder Variable berechnet (Tabelle 27).

TABELLE 27: MODELLPARAMETER DER REGRESSIONSANALYSE M9.

	Koeffizient	SF	t
<b>Intercept*</b>	3,81	0,13	29,40
<b>Winkelgröße (Bogenminute)*</b>	-0,15	0,01	-16,04
<b>Kontrast (Hoch)*</b>	-0,22	0,06	-4,02
<b>Geschwindigkeit (km/h)*</b>	-0,01	0,00	-10,66
<b>fas*</b>	-0,60	0,09	-6,75
<b>etcs</b>	-0,06	0,09	-0,66

\*  $p < 0,05$ ;  $R^2$  (Feste Effekte) = 0,44;  $R^2$  (Gesamt) = 0,56

Zufällige Effekte: SD = 0,38; ICC = 0,21

Das Tool generiert Schätzkurven für die ausgewählte Kontrast- und Zugsicherungssystem-Kombination für den Durchschnitt der in den Versuchen unter den jeweiligen Bedingungen ermittelten Reaktionszeiten. Die Schätzung der Reaktionszeit für die vom Benutzer eingegebene Bedingung wird dann auf dem Diagramm angezeigt. Ein Beispieldiagramm ist in Abbildung 34 dargestellt.

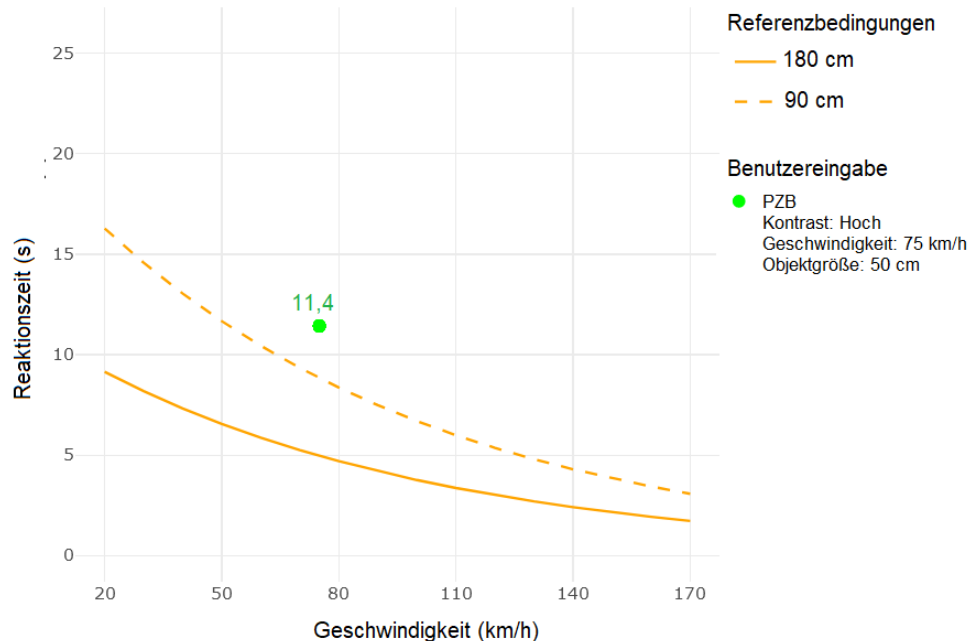


Abbildung 34: Reaktionszeitschätzung für die kontrastreichen Objekte im PZB-Szenario und die vom Benutzer eingegebene Bedingung (grüner Punkt).

Die Abbildung 34 zeigt die Reaktionszeit der Tf in Abhängigkeit von der Fahrgeschwindigkeit und der Hindernisgröße für das kontrastreiche Hindernis. Die x-Achse repräsentiert die Geschwindigkeit in km/h, während die y-Achse die Reaktionszeit in Sekunden darstellt. Zwei Referenzkurven zeigen die Reaktionszeiten für Hindernisse von 90 cm (gestrichelte Linie) und 180 cm (durchgezogene Linie). Es wird deutlich, dass die Reaktionszeit mit steigender Geschwindigkeit abnimmt und größere Hindernisse zu kürzeren Reaktionszeiten führen. Der grüne Punkt veranschaulicht eine spezifische Nutzereingabe im Tool (PZB, hoher Kontrast, 75 km/h, 50 cm Hindernis), die eine Reaktionszeit von 11,4 Sekunden ergibt. Dieses Tool bietet ein grundlegendes Beispiel für die Modellierung der Leistungsfähigkeit der Tf zur Ableitung von Benchmarkwerten.

## 11.4 Analyse der Reaktionszeit ohne Verwendung von Maßen der zentralen Tendenz

In diesem Abschnitt wurden Analysen ohne die Verwendung von Maßen der zentralen Tendenz durchgeführt. Zunächst wird die Vergleichsanalyse der Verteilungen durchgeführt, gefolgt von einer Überlebensanalyse. Der Zweck dieser Analyse besteht darin, alternative Ansätze zur Untersuchung und Interpretation der Daten zu präsentieren, die über die traditionellen zentralen Tendenzmaße hinausgehen. Durch die Betrachtung der Verteilungen können Unterschiede in den Daten aufgedeckt werden, die mit den herkömmlichen Maßen möglicherweise nicht erfasst werden. Die Überlebensanalyse ermöglicht es uns, die Zeit bis zum Eintritt eines Ereignisses zu untersuchen und die Wahrscheinlichkeit des Ereigniseintritts über die Zeit hinweg zu schätzen. Durch diese erweiterte Analyse können neue Einblicke gewonnen und möglicherweise interessante Muster und Zusammenhänge in den Daten entdeckt werden.

## 11.4.1 Vergleich der Verteilungen

In diesem Abschnitt wurde eine vergleichende Analyse der Verteilung verschiedener Bedingungen anhand ihrer Dichteverteilung und empirischen kumulativen Verteilungsfunktion (ECDF) präsentiert. Das Ziel dieser Analyse bestand darin, die Ergebnisse der linearen Regression zu ergänzen und ein tieferes Verständnis der zugrundeliegenden Muster in den Daten zu gewinnen. Durch die Untersuchung der Dichteverteilung und ECDF jeder Stufe sollen potenzielle Unterschiede und Gemeinsamkeiten zwischen den Bedingungen untersucht und die Verteilungsmerkmale der interessierenden Variablen beleuchtet werden. Dieser Ansatz ermöglicht es, die bisherigen Ergebnisse weiter zu unterstützen und zusätzliche Erkenntnisse zu gewinnen, die das Verständnis der Zusammenhänge zwischen den Variablen verbessern können.

Ein Hauptmerkmal der Reaktionszeiten ist, dass die Standardabweichung ungefähr linear mit dem Mittelwert ansteigt (Wagenmakers und Brown, 2007). Die Verschiebung nach rechts zeigt, dass die Reaktionszeiten für kleinere Objekte und langsamere Geschwindigkeiten im Bereich der längeren Reaktionszeit konzentriert sind (Abbildung 35). Eine linkssteile Dichtefunktion deutet darauf hin, dass die Reaktionszeiten im Datensatz eine Tendenz zur Konzentration auf niedrigere Werte haben. Ein längerer und dickerer Randbereich deutet auf das Vorhandensein von Ausreißern oder Extremwerten hin, die die Verteilung nach links oder rechts ziehen.

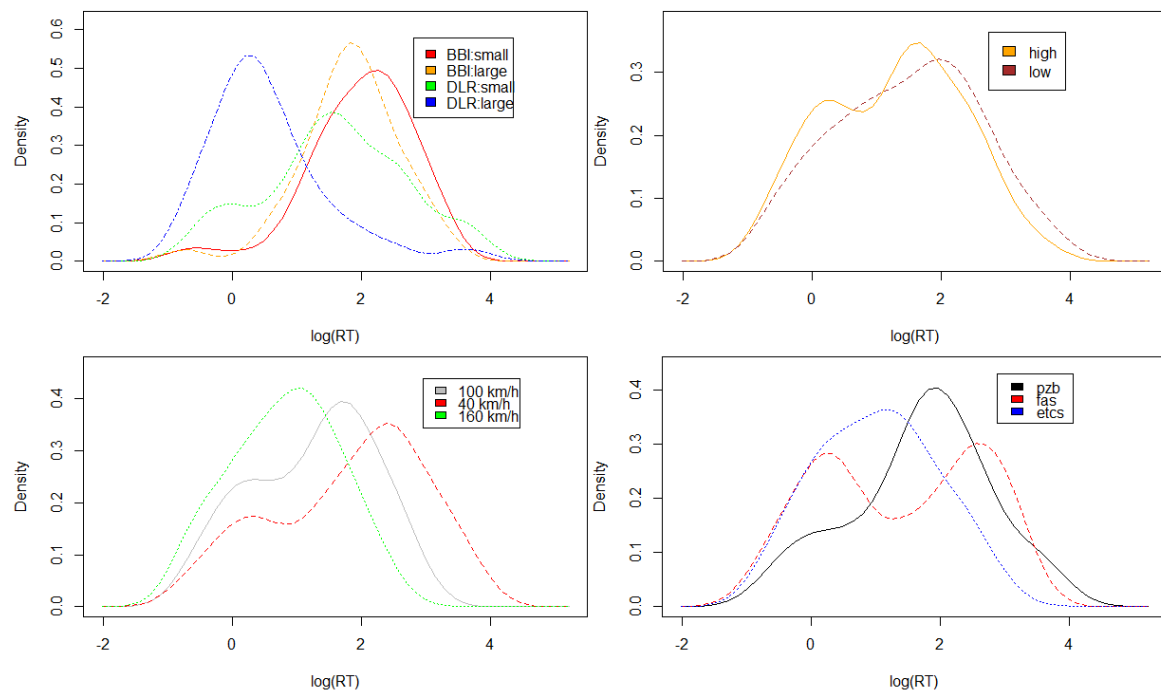


Abbildung 35: Dichtefunktionen der verschiedenen Stufen der Variablen: Objektgröße (oben-links), Kontrast (oben-rechts), Geschwindigkeit (unten-links), Zubeinflussungssystem (unten-rechts)

Doppelspitzen in der Verteilung deuten darauf hin, dass die Daten durch eine Mischung aus zwei zugrundeliegenden Verteilungen erzeugt werden können. Diese bimodale Verteilung tritt in den FaS-Daten deutlich auf. Dies könnte auf das Vorhandensein verschiedener Untergruppen innerhalb des Datensatzes hinweisen. Wie bereits erwähnt, wies das FaS-Szenario die größte Varianz zwischen den beiden Simulatorstudien auf. Die Dichtefunktionen von BBI und DLR legen nahe, dass das FaS-Szenario ähnliche Muster in Bezug auf die Reaktionszeitmessungen aufweist (Abbildung 36, links). Die Verschiebung

der Verteilung von BBI nach rechts deutet jedoch darauf hin, dass die Reaktionszeiten im Durchschnitt im BBI-Datensatz länger sind als im DLR-Datensatz. Dies ist zu erwarten, da in der DLR-Studie nur Stimuli mit großer Größe vorhanden waren, während in der BBI-Studie auch Stimuli mit kleiner Größe vorhanden waren. Dies führt dazu, dass die Reaktionszeiten in der BBI-Studie zu längeren Werten verschoben sind. Die symmetrische Form der Verteilungsschwänze legt jedoch nahe, dass es Faktoren oder Bedingungen gibt, die für jede FaS-Bedingung spezifisch sind und die die Reaktionszeiten unterschiedlich beeinflussen. Zukünftige Studien könnten vertiefende Analysen zum Fahren auf Sicht durchführen, wobei verschiedene Einflussfaktoren näher untersucht werden könnten.

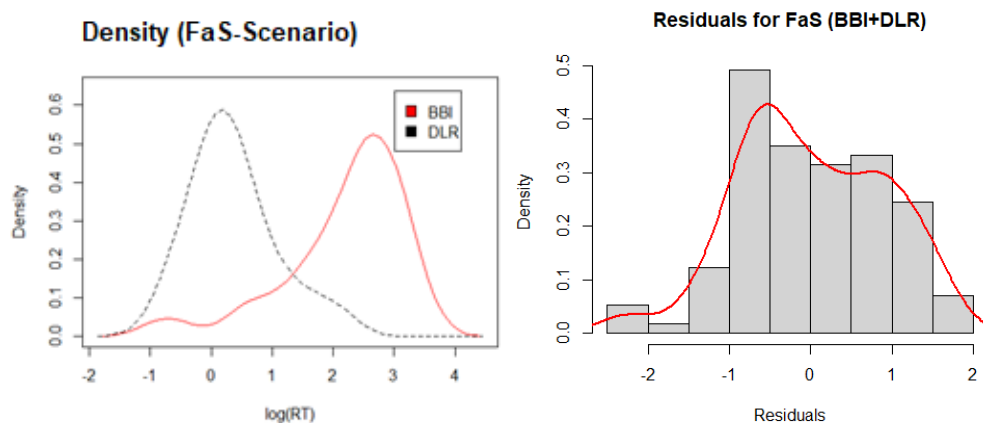


Abbildung 36: Dichteverteilung der log-transformierten Reaktionszeiten in den Studien BBI und DLR (links) und Histogramm der Residuen des gesamten Datensatzes

In der linearen Regression gilt die Annahme der Normalverteilung in der Regel für die Zufallseffekte und die Residuen, nicht jedoch für die Verteilung der unabhängigen Variablen. Die Verteilung der Residuen in der Regressionsanalyse nähert sich im Vergleich zur bimodalen Verteilung der FaS-Rohdaten einer Normalverteilung an, was auf eine bessere Anpassung hindeuten kann (Abbildung 36, rechts). Zusätzlich wurden die Residuen des gesamten Datensatzes überprüft, um die Annahmen zu bestätigen. Allerdings kann der Unterschied zwischen den beiden Verteilungen immer noch die Leistung des Modells beeinflussen. Obwohl der Zufallseffekterterm im Modell einen Teil der Varianz erklärt, müssen die zugrundeliegenden Muster im FaS-Szenario in Zukunft weiter analysiert werden.

Der geringe Unterschied zwischen den Verteilungen der Kontrastbedingungen (Abbildung 35, oben-rechts) korreliert mit den Ergebnissen der Regressionsanalyse, die darauf hindeutet, dass der Kontrast im Vergleich zu anderen Faktoren den geringsten signifikanten Einfluss auf die Reaktionszeit hat.

Die Reaktionszeitverteilungen zwischen den Gruppen wurden sowohl durch visuelle Inspektion der empirischen kumulativen Verteilungsfunktion (ECDF) als auch durch gepaarte Zweistichproben-Kolmogorov-Smirnov-Tests verglichen. Dies ermöglichte es, zu untersuchen, ob die Daten innerhalb der beiden Gruppen aus derselben Wahrscheinlichkeitsverteilung stammen. Die x-Achse eines ECDF-Plots repräsentiert die logarithmierte Reaktionszeit, und die y-Achse repräsentiert den Anteil der Datenpunkte, die einen Wert kleiner oder gleich einem gegebenen Wert auf der x-Achse haben (Abbildung 37).

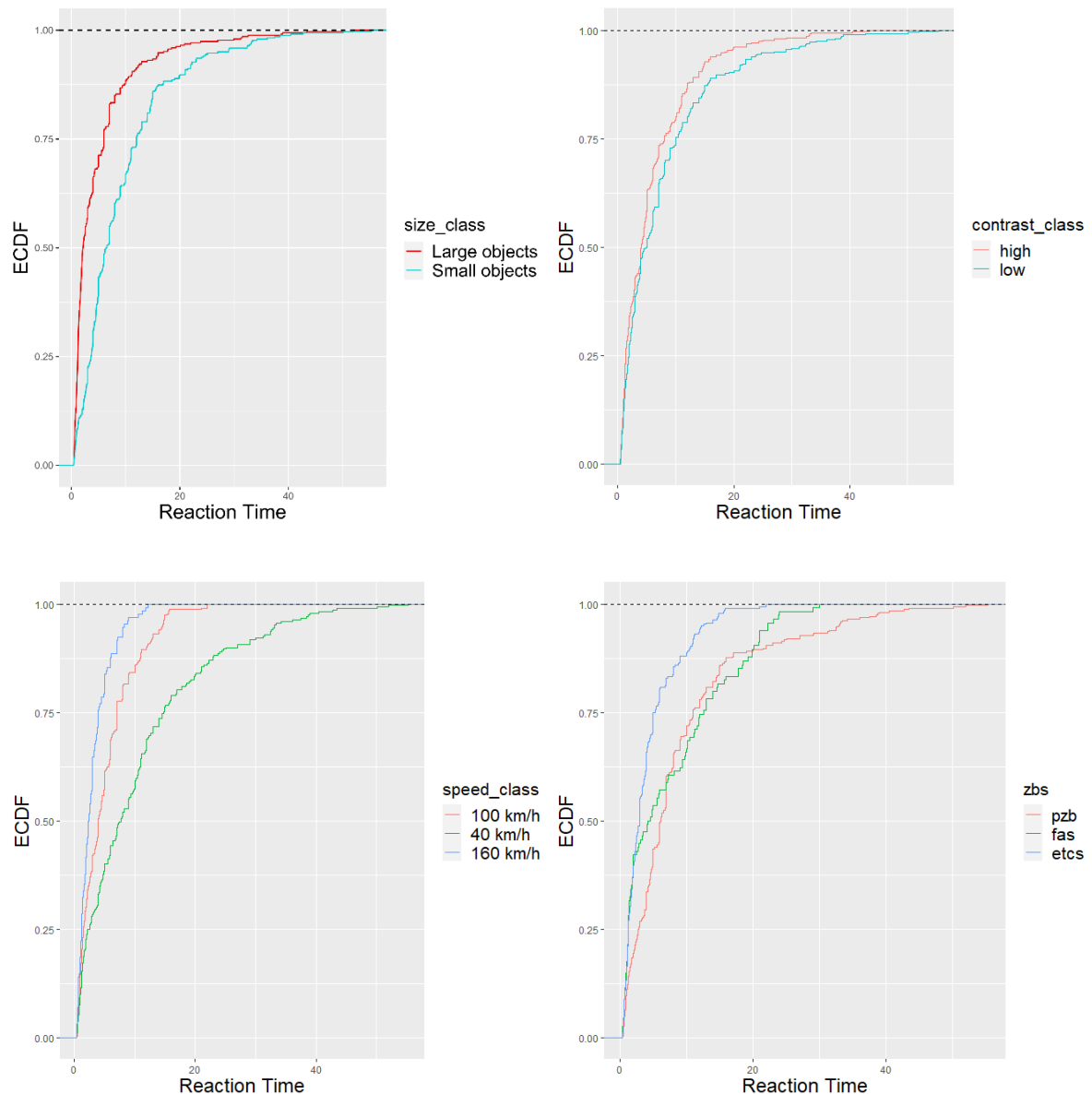


Abbildung 37: Empirische kumulative Verteilungsfunktionen der verschiedenen Stufen der Variablen: Objektgröße (oben-links), Kontrast (oben-rechts), Geschwindigkeit (unten-links), Zugsicherungssystem (unten-rechts)

Die Kurven repräsentieren die kumulative Wahrscheinlichkeit, dass ein Wert in einem Datensatz kleiner oder gleich einem bestimmten gegebenen Wert ist. Wenn eine Kurve durchweg über den anderen Kurven liegt, deutet dies darauf hin, dass die entsprechende Bedingung oder Gruppe tendenziell niedrigere Werte aufweist. Dies ist in den Plots mit unterschiedlichen Größen- und Geschwindigkeitsstufen zu sehen (Abbildung 37, links). Obwohl die Kurve der hohen Kontrastbedingung durchgehend über der Kurve der niedrigen Kontrastbedingung liegt, scheint der Unterschied sehr gering zu sein (Abbildung 37, oben-rechts). Eine steile Kurve am Anfang deutet darauf hin, dass die Wahrscheinlichkeit, kürzere Reaktionszeiten zu beobachten, größer ist. Im Gegensatz dazu zeigt eine flachere Kurve, dass die Konzentration um bestimmte Werte herum geringer ist. Beispielsweise hat die Kurve der Reaktionszeiten für große Objekte relativ steile Abschnitte zwischen den Werten 0 und 10, was darauf hinweist, dass die Wahrscheinlichkeit von Reaktionszeiten in diesem Bereich sehr hoch ist. Im Gegensatz dazu ist die Reaktionszeit ab 15 Sekunden beim PZB-Szenario relativ flach, was darauf hindeutet, dass die Wahrscheinlichkeit

höherer Reaktionszeiten aufgrund weniger Datenpunkte in einem breiten Bereich gering ist (Abbildung 37, unten links). Der Plot der Zugsicherungssysteme deutet darauf hin, dass es Unterschiede in den Verteilungen der ETCS-FaS-Datensätze für höhere Werte gibt. Für beide Datensätze liegen rund 45 % der Reaktionszeitwerte unter 3 Sekunden (Abbildung 37, unten-rechts). Ab diesem Punkt verteilt sich die Reaktionszeitverteilung von FaS stärker. Auf der anderen Seite wird die Kurve von PZB am Ende weniger steil, was auf mehrere Ausreißer im Bereich langer Reaktionszeiten hindeuten könnte. Während die visuelle Inspektion der Größen- und Geschwindigkeitsplots klare Unterschiede zwischen den getesteten Verteilungen zeigt, ist ein statistischer Test für die Kontrast- und ZSS-Plots erforderlich.

Der KS-Test prüft, ob ein signifikanter Unterschied zwischen den Verteilungen zweier gepaarter Stichproben besteht. Dazu wird die maximale Differenz zwischen den empirischen kumulativen Verteilungsfunktionen (ECDFs) der beiden Stichproben verglichen. Die Teststatistik ist die maximale absolute Differenz (D-Statistik) zwischen den beiden ECDFs. Die Nullhypothese für den gepaarten Kolmogorov-Smirnov-Test mit zwei Stichproben lautet, dass die beiden Stichproben aus der gleichen Verteilung stammen. Die P-Werte des KS-Tests deuten darauf hin, dass alle Paare bei einem Konfidenzniveau von 0,05 einen signifikanten Unterschied aufweisen (Tabelle 28).

TABELLE 28: KOLMOGOROW-SMIRNOW-TEST-PARAMETER FÜR VERGLICHENE PAARE

Paar	Kolmogorov-Smirnov-Test		
		D	p
klein - groß		0,389	< 0,01
kontrastarm - kontrastreich		0,116	0,02
40 km/h - 100 km/h		0,294	< 0,01
100 km/h - 160 km/h		0,261	< 0,01
FaS - PZB		0,229	< 0,01
PZB - ETCS		0,335	< 0,01

Durch Überlagerung der ECDF der erhobenen Daten und der durch das Regressionsmodell vorhergesagten Werte kann man bewerten, wie gut das Modell die Verteilung der Daten abbildet (Abbildung 38). Die Vorhersagen stellen die erwarteten Werte dar, die sich aus den Koeffizienten des Regressionsmodells in Abschnitt 11.3.2 ergeben. Eine etwas steilere Kurve der Vorhersagen im Vergleich zu den Beobachtungen deutet darauf hin, dass das Modell einen engeren Bereich der Reaktionszeiten abdeckt. Dies ist vor allem auf die geringere Abdeckung der Extremwerte durch das Modell zurückzuführen. Der Kolmogorov-Smirnov-Test zeigte jedoch, dass sich die Verteilung der erhobenen und vorhergesagten Werte nicht signifikant unterscheidet ( $D = 0,218$ ,  $p = 0,076$ ).

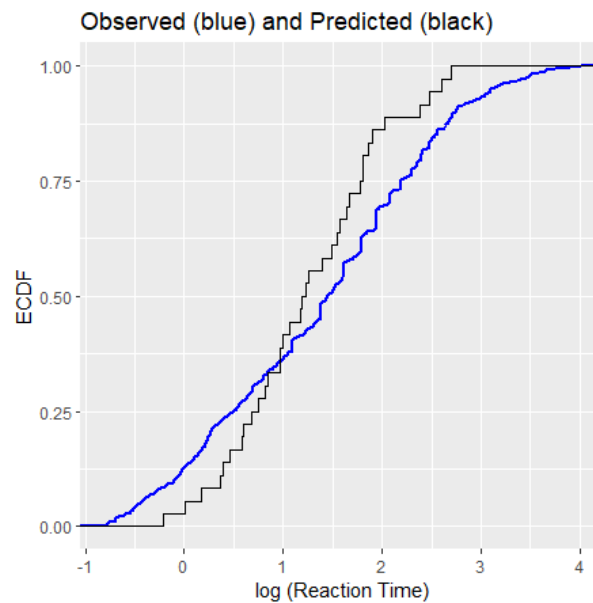


Abbildung 38: Empirische kumulative Verteilungsfunktion (ECDF) der beobachteten Werte und vorhergesagten Werte für den gesamten Datensatz

## 11.4.2 Probabilistisches Modell

Nach der Beschreibung der empirischen Verteilung der Beobachtungen anhand der ECDF-Daten wurden die Daten in diesem Abschnitt an ein probabilistisches Modell angepasst, um zu bewerten, wie verschiedene Faktoren die Reaktionszeit der Probanden beeinflussen.

Ereigniszeitanalyse (eng. Survival Analysis) ist ein probabilistischer Modellierungsansatz zur Analyse von Zeit-zu-Ereignis-Daten, um die Wahrscheinlichkeitsverteilung der Zeit bis zum Eintreten eines Ereignisses zu modellieren (Kleinbaum und Klein, 2012). Im Bereich der Verkehrsforschung umfassen Beispiele die Zeit bis zum Auftreten eines Fahrzeugunfalls und die Zeit bis zur Einführung neuer Verkehrstechnologien (Washington, 2020). Es gibt mehrere Studien, die das parametrische Ereigniszeitanalyse mit Accelerated Failure Time (AFT) anwenden, um die Auswirkungen von Kovariaten zu analysieren. In mehreren Studien wurden die Reaktionszeiten von Autofahrern auf Stimuli (z. B. Fußgänger) mit der Nutzung von Mobiltelefonen als Ablenkungskovariate modelliert (Haque und Washington, 2014, Choudhary und Velaga, 2017, Liu, Alsaleh und Sayed, 2021), während Pawar und Velaga (2020) die Auswirkung von Zeitdruck auf die Reaktionszeiten anhand einer Ereigniszeitanalyse untersuchte. Im Eisenbahnbereich sind Ereigniszeitanalyse in den Bereichen der präventiven Instandhaltung zu finden (Andersson, Björklund und Haraldsson, 2016).

Bei dem AFT-Modell skaliert der Effekt von Kovariaten (d. h. unabhängige Variablen) die Dauervariable (z. B. Reaktionszeit) direkt in der Basisüberlebensfunktion und liefert bessere Ergebnisse, wenn mehr als eine Kovariate vorhanden ist (Choudhary und Velaga, 2017). Für ein parametrisches AFT-Modell ist auch der Typ der Verteilung für die Dauervariable erforderlich. Die Weibull-Verteilung ermöglicht Veränderungen der Wahrscheinlichkeit, dass ein Ereignis im Laufe der Zeit eintritt (Washington, 2020). Positive Dauerabhängigkeit bedeutet, dass die Wahrscheinlichkeit des Ereignisses im Laufe der Zeit zunimmt ( $P > 1$ ). Darüber hinaus ermöglicht die Weibull-Verteilung eine einfache Interpretation.

Es wurde ein Weibull-AFT-Modell verwendet, um die Beziehung zwischen der Reaktionszeit und den erklärenden Variablen zu untersuchen. In dem Modell wird das Ereignis als die Erkennung des Objekts

durch die Fahrer definiert. Die Dauer wird als die Zeit bis zur Reaktion der Teilnehmer auf die Objekte definiert. Das Modell wurde mit dem vollständigen Datensatz ohne zufällige Effekte mithilfe der survival- und flexsurv-Pakete in R (Jackson, 2016) angepasst.

Dichtefunktion des Weibull-Modells, wobei  $\lambda$  und  $P$  die Standort- und Formparameter der Weibull-Verteilung sind:

$$f(t) = P\lambda(\lambda t)^{P-1} \exp(-(\lambda t)^P)$$

Die Hazardfunktion und die Überlebensfunktion des Modells werden wie folgt ausgedrückt:

$$H(t) = (\lambda P)(\lambda t)^{P-1}$$

$$S(t) = \exp[-(\lambda t)^P]$$

Der Einfluss externer Kovariaten auf die Überlebenszeit kann wie folgt formuliert werden, wobei  $X$  eine Kovariate innerhalb des Vektors der Kovariaten ( $X_i$ ) und  $\beta_{0-n}$  die Koeffizienten darstellt:

$$\lambda = \exp([-P(\beta_0 + \beta_1 X_0 + \dots)])$$

Aufgrund des Versuchsdesigns wurde die Erkennung jedes Objekts durch die Teilnehmer erfasst. Zensierte Daten beziehen sich auf Beobachtungen, bei denen das interessierende Ereignis (in diesem Fall die Objekterkennung) noch nicht eingetreten ist oder innerhalb des Untersuchungszeitraums nicht vollständig beobachtet werden konnte (Kleinbaum & Klein, 2020). Solche zensierten Überlebenszeiten unterschätzen die wahre (aber unbekannte) Zeit bis zum Ereignis (Clark et al., 2003). Ohne zensierte Daten liegen vollständige Informationen über das Auftreten von Ereignissen für alle Teilnehmenden vor. Daher liefern die Daten mehr Informationen über die Form der Verteilung (Klein und Moeschberger, 2003).

### 11.4.3 Ergebnisse der Überlebensanalyse

Die Modellparameter sind in Tabelle 29 aufgeführt. Der geschätzte Formparameter  $P$  für die Weibull-Verteilung beträgt 1.326, was darauf hinweist, dass die Wahrscheinlichkeit des Ereignisses (Objekterkennung) mit zunehmender Zeit steigt. Der Skalenparameter spiegelt die Zeitskala des Überlebensergebnisses wider und gibt den Zeitpunkt an, zu dem das Ereignis am wahrscheinlichsten eintritt. Der Unterschied in den Log-Likelihoods und der signifikante  $p$ -Wert des Chi-Quadrat-Tests deuten darauf hin, dass das Modell mit Kovariaten besser zu den Daten passt als das Modell nur mit dem Intercept. Die Referenzebene des Modells ist wie folgt: PZB, 100 km/h, 3,78° Winkelgröße des Objekts, hoher Kontrast. Die negativen Schätzwerte deuten auf eine Abnahme der Reaktionszeiten im Vergleich zur Referenzebene hin. Zum Beispiel erhöht sich die Reaktionszeit auf Objekte mit geringem Kontrast um 32 % im Vergleich zu Objekten mit hohem Kontrast., Eine Erhöhung der Geschwindigkeit auf 160 km/h ist mit einer Abnahme der Reaktionszeiten um 37 % verbunden.

TABELLE 29: SCHÄTZUNGEN DES WEIBULL-AFT-MODELLS

Variable	Koeffizient	Exp( $\beta$ )	KI 5 %	KI 95 %	SF
<b>Winkelgröße_6,72*</b>	-0,23	0,80	-0,39	-0,06	0,09
<b>Winkelgröße_7,56*</b>	-0,33	0,72	-0,52	-0,14	0,10
<b>Winkelgröße_13,43*</b>	-1,22	0,30	-1,38	-1,07	0,08
<b>Geschw._40km/h*</b>	1,16	3,19	0,98	1,34	0,09

Variable	Koeffizient	Exp( $\beta$ )	KI 5 %	KI 95 %	SF
Geschw._160km/h*	-0,46	0,63	-0,46	-0,64	0,09
Kontrast_niedrig*	0,28	1,32	0,28	0,16	0,06
zss_fas*	-0,75	0,47	-0,95	-0,55	0,10
zss_etcs	0,01	1,01	-0,16	0,19	0,10
<b>Goodness-of-fit statistics</b>					
Log-likelihood (Model)	-1813	Shape (P)	1,32		
Log-likelihood (intercept-only)	-2063	Scale ( $\lambda$ )	6,78		
Chsiq (df)	432* (10)	AIC	3646		
Anzahl der Iterationen	5	*p<0,05			
Anzahl der Beobachtungen	609				

Mit Hilfe von Hazard- und Überlebensfunktionen lassen sich Median- (oder andere Quantil-)Reaktionszeiten und Wahrscheinlichkeiten für bestimmte Bedingungen berechnen. Zum Beispiel beträgt zum Zeitpunkt  $t = 5$  Sekunden die Wahrscheinlichkeit, dass das Objekt im PZB-Szenario nicht erkannt wird unter den Basiskonditionen ( $S_5 | pzb$ ) etwa 51 % (Abbildung 39). Um diesen Wert für das größte Objekt (13,43') zu berechnen, müssen die entsprechenden Variablen in der Überlebensfunktion ( $S(t)$ ) angepasst werden. In diesem Fall sind die Variable  $X_{13,43zbs}$  auf 1 gesetzt, wobei alle anderen Variablen auf ihre Basiskonditionen gesetzt wurden. Das Ergebnis beträgt 35 % (Abbildung 39) für das große und kontrastreiche Objekt bei 100 km/h.

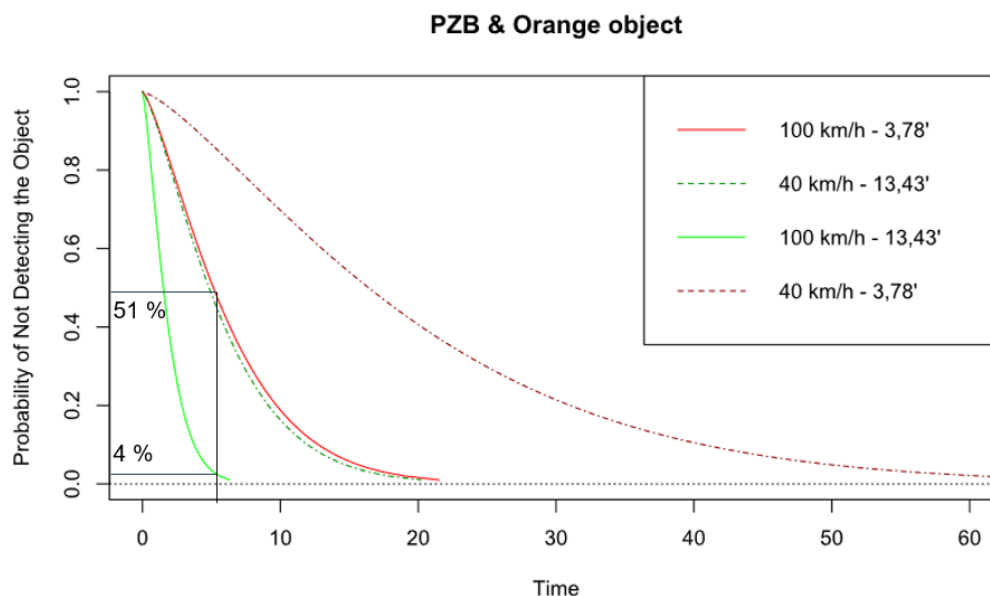


Abbildung 39: Wahrscheinlichkeiten der Nichterkennung von Objekten für verschiedene Bedingungen. Die X-Achse ist die Zeit seit dem Erscheinen des Objekts.

Vergleicht man FaS und PZB für die beiden Objekte der DLR-Studie bei einer Geschwindigkeit von 40 km/h, so ist die Wahrscheinlichkeit, das Objekt zu erkennen, unabhängig von den Kontraststufen bei PZB zu jedem Zeitpunkt höher (Abbildung 40). Die Überlebensanalyse liefert wertvolle Erkenntnisse über die Wahrscheinlichkeit, dass ein Ereignis (in diesem Fall die Objekterkennung) über einen bestimmten Zeitraum eintritt. Wie in Abbildung 39 und Abbildung 40 dargestellt, kann man auch die Unterschiede in den Wahrscheinlichkeiten zwischen verschiedenen Szenarien oder Niveaus der erklärenden Variablen hervorheben und so den Einfluss jedes Faktors auf das Ereignis von Interesse betonen.

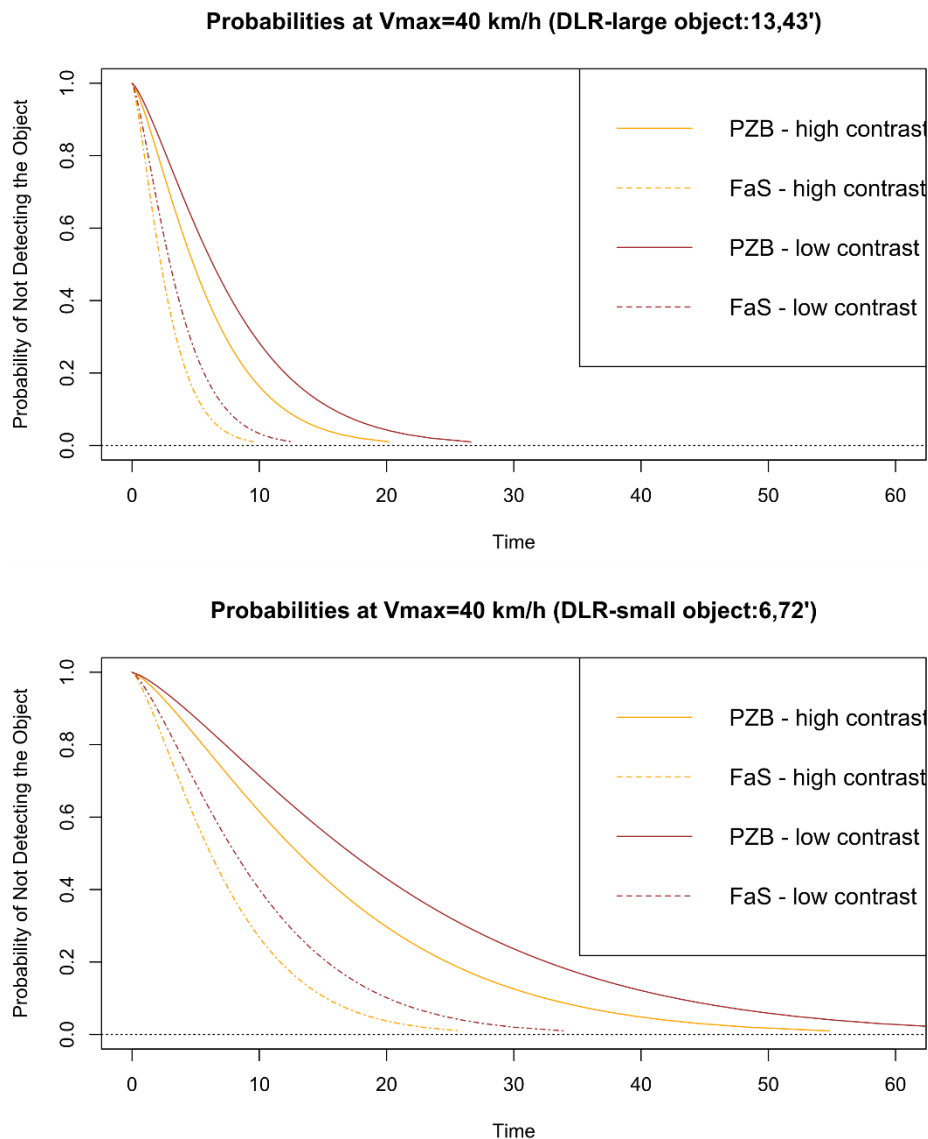


Abbildung 40: Wahrscheinlichkeiten, die Objekte bei 40 km/h nicht zu erkennen.

## 11.5 Zusammenfassung der Ergebnisse

Die Ergebnisse der Regressionsanalyse und der Analysen ohne zentrale Tendenzen lieferten Einblicke in die Effekte der unabhängigen Variablen auf die Reaktionszeit (RT). Die Ergebnisse wurden sowohl graphisch als auch statistisch präsentiert, was eine detaillierte Interpretation der Effekte ermöglichte. Zusätzlich wurde eine Überlebensanalyse verwendet, um das Verhalten der RT in Bezug auf das Ereignis der Objekterkennung zu untersuchen.

Die Analyse der festen Effekte im Mixed-Effekte-Modell zeigte signifikante Effekte von Größe, Kontrast und Geschwindigkeitstyp auf die Reaktionszeiten. Die Koeffizienten von Größe, Kontrast und Geschwindigkeitstyp gingen in beiden Simulatoren in die gleiche Richtung und hatten einen ähnlichen relativen Einfluss. Dabei ist der Einfluss der Größe mindestens doppelt so hoch wie der Einfluss von Kontrast. Der

Einfluss der Geschwindigkeit von 40 vs. 100 km/h war in beiden Fällen doppelt so groß wie die Geschwindigkeit von 100 vs. 160 km/h. Die Hypothese, dass das Fahren mit ETCS im Vergleich zum Fahren mit PZB zu höheren Reaktionszeiten führt, konnte nicht bestätigt werden. Ähnliche Zusammenhänge wurden auch bei der Analyse des gesamten Datensatzes festgestellt.

Die Analyse der zufälligen Effekte zeigte, dass die Simulationsgruppe einen signifikanten Einfluss auf die RT hatte, was auf eine Variabilität der RT zwischen verschiedenen Simulatoren hinweist. Das liegt vor allem an der unterschiedlichen Bildschirmgröße und Auflösung von BBI- und DLR-Simulatoren. Es wurde nachgewiesen, dass ein großer Teil der Varianz zwischen den Simulatoren auf die Unterschiede in der wahrgenommenen Größe der Stimuli zurückzuführen ist. Die Modellanpassung wurde verbessert, indem ein Teil der Varianz zwischen den Simulatoren durch Unterschiede in der Objektgröße als fester Effekt erfasst wurde.

Insgesamt unterstützte die visuelle und statistische Untersuchung der Verteilungen der Reaktionszeit die Ergebnisse der Regressionsanalyse. Durch die Untersuchung der Verteilungen konnten Unterschiede aufgedeckt werden, die durch herkömmliche Maße möglicherweise nicht erfasst wurden. Es wurde festgestellt, dass die lineare Regressionsanalyse nicht genügend Beweise für statistische Unterschiede zwischen den PZB- und ETCS-Bedingungen liefert (Satterthwaite's  $p > 0,05$ ). Beim Vergleich der kumulativen Verteilungsfunktionen dieser Datensätze ergab der Kolmogorov-Smirnov-Test jedoch einen p-Wert von weniger als 0,05, was darauf hindeutet, dass die Datensätze nicht aus der gleichen Verteilung stammen. Diese Diskrepanz deutet darauf hin, dass, obwohl sich die Mittelwerte oder andere durch die lineare Regressionsanalyse geschätzte Parameter zwischen den PZB- und ETCS-Bedingungen nicht signifikant unterscheiden, es Unterschiede in den Gesamtverteilungsformen oder -mustern gibt. Die komplexe Beziehung zwischen Reaktionszeit und Zugsicherungssystemen sowie anderen zugrundeliegenden Faktoren muss in zukünftigen Forschungen weiter untersucht werden.

Die Überlebensanalyse lieferte weiteres Verständnis, indem sie die Zeit bis zum Eintritt eines Ereignisses untersuchte und die Wahrscheinlichkeit des Ereigniseintritts im Laufe der Zeit schätzte. Die Diagramme zeigen die Wahrscheinlichkeit, dass die Ereignisse nicht eintreten (d. h. das Erkennen des Objekts), als Funktion der Zeit. Diese Dauer berücksichtigt Kovariaten, die den Eintritt der Objekterkennung beeinflussen können. Die Ergebnisse der Überlebensanalyse ergänzen die Ergebnisse anderer analytischer Ansätze wie der linearen Regression und Verteilungsanalysen und tragen zu einer umfassenden Bewertung der Daten bei. Ausreißer sind Datenpunkte, die signifikant von der Mehrheit der Daten abweichen und einen starken Einfluss auf statistische Analysen wie die Regressionsanalyse haben können.

Die Ergebnisse dieser Studie tragen zur Erweiterung des Wissens über die Faktoren bei, die die RT beeinflussen, und bieten eine Grundlage für die Vorhersage von RT unter verschiedenen Bedingungen. Es bleibt jedoch anzumerken, dass weitere Forschung erforderlich ist, um die genauen Mechanismen hinter diesen Beziehungen zu verstehen und mögliche Einflussfaktoren wie Aufmerksamkeit und andere nicht berücksichtigte Variablen zu ermitteln. Eine größere Anzahl von Daten, andere Messungen wie die Blickrichtung und ein besser ausgewogenes Versuchsdesign könnten zuverlässigere Ergebnisse liefern.

## 11.6 Limitationen

Ziel der Simulatorstudien war es, die Wahrnehmungsleistung der visuellen Wahrnehmung in verschiedenen Situationen zu messen. Die Bedingungen, unter denen teilnehmende Tf die ausgewählte Aufgabe ausführten, wurden durch die Verwendung von unabhängigen Variablen manipuliert. Wie die

Wahrnehmung im menschlichen Auge funktioniert oder die Informationen von Tf kognitiv verarbeitet werden, ist nicht Gegenstand dieses Projekts.

Es ist zu beachten, dass es sich um eine simulatorbasierte Studie handelt. Fahr simulatoren haben zwar viele Vorteile wie einfache Datenerfassung, Reproduzierbarkeit und Kontrollierbarkeit, bringen aber auch Nachteile mit sich. Die Arbeit eines Tf ist hochkomplex und die visuelle Überwachung des Bereichs entlang der Gleise wird durch ein kompliziertes Zusammenspiel von Faktoren beeinflusst, die in den Studien nicht alle berücksichtigt werden konnten. Die simulierte Fahrsituation unterscheidet sich aufgrund zahlreicher Grenzen bei der Simulation der virtuellen Umgebung oft stark von der realen Fahrsituation. Die Leuchtdichte, die Blendung und der Kontrast auf den Bildschirmen des Simulators können sich von den Bedingungen in der Realität unterscheiden. Zudem könnte die Motivation die Aufmerksamkeit der Teilnehmenden in Simulatorstudien erhöhen. Andererseits könnte das Fehlen eines Risikos durch gefährliche Ereignisse in einem Simulator die Aufmerksamkeit verringern.

Eine weitere Begrenzung besteht darin, dass die Teilnehmenden die Anweisung erhalten hatten, auf Objekte zu reagieren, die sie erkennen. Die potenziellen Auswirkungen der Erwartung auf die visuelle Wahrnehmung wurden nicht berücksichtigt. Die Wahrnehmungsprozesse der visuellen Detektion und Erkennung könnten von probabilistischen Informationen über das wahrscheinliche Auftreten von Objekten durch assoziatives Lernen, Priming oder kontextbezogene Verarbeitung von Szeneninformationen profitieren (Summerfield und Egner, 2009). Da die Objekterkennung jedoch eine übliche Aufgabe der Tf ist, ist diese Anweisung nicht völlig inkompatibel mit der Praxis des Bahnbetriebs. Um die Auswirkungen der Erwartungshaltung auf die sensorische Leistungsfähigkeit in den Experimenten zu begrenzen, wurde versucht, die räumliche Ungewissheit zu erhöhen, d. h. die Objekte tauchten nicht an Orten mit hoher Erwartungshaltung auf.

Gemäß der Ril 408.2301 Abschnitt 1 müssen Tf in Deutschland für jede der von ihnen befahrenen Bahnstrecke streckenkundig sein (Richtlinie 408.21-27:2014). Ohne Streckenkenntnis dürfen Tf eine Strecke nur nach den in der Richtlinie angegebenen Zusatzregeln fahren. Diese wird für die eigenverantwortliche, sichere und fahrplanmäßige Befahrung der Strecke benötigt. So könnten beispielsweise Tf mit Ortskenntnis bei der Annäherung an einen Bahnübergang potenzielle Gefahren besser vorhersehen. Andererseits könnte diese Erwartungshaltung zu einer falschen Interpretation von Informationen und zu im Voraus geplanten Verhaltensweisen führen, die unangemessen sind (McLeod, Walker und Moray, 2005). Selbst wenn vor dem Experiment ein Mittel zum Erwerb von Streckenkenntnis zur Verfügung gestellt wurde, ist es schwierig, seinen Einfluss auf jeden Einzelnen zu messen, um seine Wirkung auf die sensorische Gesamtleistung zuverlässig zu beurteilen.

## 12 Diskussion

Für eine ganzheitliche Einordnung der Projektergebnisse in den Kontext der Entwicklung von Anforderungen an ein zukünftiges ATO-System ist es zunächst wichtig, die Übertragbarkeit der Studienergebnisse auf den realen Bahnbetrieb zu analysieren. Dabei werden im Folgenden verschiedene Faktoren dargestellt, die die Übertragbarkeit beeinflussen. Die Faktoren sind hier genau wie die unabhängigen Variablen nach persönlichen Einflussfaktoren, betrieblichen Parametern und physikalischen Eigenschaften des Reizes sortiert. In Bezug auf die persönlichen Einflussfaktoren hat die Situation der Erhebung im Rahmen einer Simulatorstudie sicherlich Auswirkungen auf die Ergebnisse. Durch die Instruktionen waren die Teilnehmenden informiert, dass sie einen Reiz zu erwarten hatten, auf den sie reagieren sollten. Demzufolge war vermutlich die visuelle Suche in der Umgebung besonders ausgeprägt, vielleicht stärker als bei einer realen Fahrt. Auf der anderen Seite steht zu bedenken, dass bei einer realen Fahrt eine der zentralen Aufgaben des Tf ist, den Gleisbereich zu überwachen. Auch bei einer realen Fahrt sollte also eine aufmerksame visuelle Überwachung der Infrastruktur stattfinden. Zudem besteht im Rahmen einer Simulatorfahrt bei Unaufmerksamkeit keine echte Gefahr, was wiederum die empfundene Dringlichkeit der visuellen Suche reduzieren könnte.

Des Weiteren wurde im Rahmen der Studien zunächst nur eine einfache Reaktionsaufgabe untersucht. Die Teilnehmenden sollten auf jeden der identifizierten Reize reagieren, es war nicht notwendig, eine Entscheidung darüber zu treffen, ob das entdeckte Objekt eine Gefahr darstellt. Im realen Betrieb wird ein Tf nicht auf jedes wahrgenommene Objekt auf die gleiche Weise reagieren. Je nach Objekt kann z. B. eine Notbremsung, ein Betätigen des Signalhorns oder auch keine Reaktion erfolgen. Dies bedeutet, dass sich die hier erhobenen Reaktionszeiten im realen Betrieb vermutlich verlängern würden. Zusätzlich zu der Zeit, die es benötigt, ein Objekt wahrzunehmen und eine physische Reaktion einzuleiten, würde dann im realen Betrieb ein Zeitraum zur Verarbeitung der wahrgenommenen Informationen sowie zur Entscheidung über eine angemessene Reaktion hinzukommen.

Weitere Punkte bezüglich der persönlichen Einflussfaktoren sind die Dauer der Fahrtzeit sowie die fehlende Streckenkenntnis. Durch den relativ kurzen Fahrtzeitraum in den Studien wurden negative Effekte durch Müdigkeit oder Abnahmen in der Vigilanz (Kapitel 5) begrenzt. Dies bedeutet, dass unter realen Betriebsbedingungen, z. B. nach einer siebenstündigen Fahrt, die Reaktionszeiten bei der visuellen Wahrnehmung schlechter ausfallen könnten als in den Studien erhoben. Auf der anderen Seite hatten die Teilnehmenden keine Streckenkenntnis über die im Simulator gefahrenen Strecken. Dies kann dazu geführt haben, dass sie ineffektive visuelle Suchstrategien anwendeten. Haben Tf Streckenkenntnis, so ist zu erwarten, dass sie genau wissen, an welchen Stellen der Strecke mit höherer Wahrscheinlichkeit wahrzunehmende Objekte auftauchen könnten (z. B. Bahnübergänge). Dies würde die Reaktionszeiten im Falle eines Objekts an einer erwarteten Position positiv beeinflussen.

Bezüglich der in der Simulatorstudie umgesetzten betrieblichen Randbedingungen lassen sich in Bezug auf die Übertragbarkeit vor allem die Komplexität der betrachteten Aufgabe sowie die Streckengeometrie betrachten. Erstens: Die teilnehmenden Tf hatten in den Studien lediglich die Aufgabe, ihren Zug zu fahren. Im realen Betrieb kommen aber weitere Aufgaben wie z. B. Kommunikation mit den Fahrdienstleitenden oder den Passagieren hinzu, Ablesen weiterer Instrumente wie Buch- und Langsamfahrplan oder Berücksichtigung des Diagnose-/Fehlerdisplays entfielen. Es stellt sich die Frage, wie diese zusätzlichen Aufgaben die visuelle Wahrnehmung eines Tf im realen Betrieb beeinflussen. Theoriegeleitet sollten auditive/verbale Aufgaben wie die Kommunikation die visuelle Leistungsfähigkeit nicht negativ beeinflussen (Wickens, 2002)). Auf der anderen Seite wären durch weitere visuelle Aufgaben wie die Überwachung des Fehlerdisplays oder Fahrplans negative Effekte auf die Leistungsfähigkeit der visuellen

Wahrnehmung bei der Überwachung der Infrastruktur zu erwarten. Zweitens wurden die wahrzunehmenden Objekte in den durchgeführten Studien immer unter idealen Sichtbedingungen positioniert - an geraden Strecken mit geringer oder keiner Steigung, ohne verdeckende Vegetation o. ä. Demzufolge war es für die Teilnehmenden möglich, die Objekte auf die Entfernung von 800 m wahrzunehmen. Unter realen Betriebsbedingungen ist die Sicht der Tf jedoch häufig stärker eingeschränkt, sei es durch Kurven, Steigungen oder Bewuchs entlang der Strecke.

Abschließend war die Variation der Eigenschaften des in den Studien wahrzunehmenden Objekts begrenzt. Es wurde ein stationärer Würfel betrachtet. Dementgegen stehen die Objekte, die ein Tf im realen Betrieb wahrzunehmen hat. Diese sind häufig nicht stationär, haben unterschiedliche Formen und Farben. Dies hat einen Einfluss auf die Wahrnehmbarkeit der Objekte. Zusammenfassend konnten im Rahmen der Studien Ergebnisse zu einem klar abgegrenzten Bereich in der visuellen Wahrnehmung erhoben werden. Darüber hinaus wäre es für ein umfassendes Bild der Leistungsfähigkeit der visuellen Wahrnehmung von Tf sinnvoll, weitere Parameter wie z. B. komplexere Aufgaben, andere wahrzunehmende Objekte oder längere Fahrtzeiten auf bekannten Strecken zu untersuchen.

Aufbauend auf den Studienergebnissen gäbe es zudem verschiedenen Möglichkeiten, diese Ergebnisse in weiteren Studien zu validieren. Die erste Möglichkeit ist die Durchführung von weiteren Replikationsstudien im Fahrsimulator. Im vorliegenden Projekt wurden bereits zwei Simulatoren eingesetzt, eine Vergleichbarkeit der Ergebnisse über unterschiedliche Simulatoren sowie zwei Stichproben konnte demnach gezeigt werden. Die Durchführung weiterer Replikationen würde die Validität der Ergebnisse weiter belegen. Zweitens wäre die Durchführung von Realfahrtstudien gerade im Hinblick auf eine Validierung der Übertragbarkeit der Studienergebnisse in die reale Welt denkbar. Eine Umsetzung von Realfahrtstudien, in denen der hier untersuchte Versuchsaufbau wiederholt werden würde, wäre jedoch äußerst komplex.

Zusätzlich zu den oben genannten Faktoren trägt der psychologische und kognitive Zustand der Tf während des realen Betriebs zu einer weiteren Komplexität bei, die die Übertragbarkeit der Studienergebnisse erschwert. Die kognitive Belastung und das Stressniveau, die durch Faktoren wie Arbeitsbelastung, zwischenmenschliche Dynamik und Umweltstressoren beeinflusst werden können, könnten die visuelle Wahrnehmung und die Entscheidungsfähigkeit der Fahrer erheblich verändern. Während eine Simulatorstudie diese Variablen kontrollieren kann, um eine einheitliche Umgebung für alle Teilnehmenden zu schaffen, führt die dynamische und oft unvorhersehbare Natur des realen Zugbetriebs ein Maß an Variabilität ein, das in einer kontrollierten Umgebung nur schwer zu reproduzieren ist. Diese Variabilität könnte zu Abweichungen bei den Leistungsmetriken der visuellen Wahrnehmung und der Reaktionszeit im Vergleich zu den in den Simulatorstudien beobachteten Werten führen. Um dem Rechnung zu tragen, sollten künftige Bewertungen von ATO-Systemen realistische Simulationen der kognitiven Belastung und des Stresses einbeziehen oder idealerweise Feldstudien durchführen, die das gesamte Spektrum der Betriebsbedingungen abbilden, denen Tf begegnen können.

Abschließend könnte es insbesondere im Hinblick auf die Reaktionszeit bei der visuellen Wahrnehmung unter realen Betriebsbedingungen sinnvoll sein, sich näher mit dem Entscheidungsfindungsprozess der Tf zu beschäftigen, der nach der Objektwahrnehmung stattfindet. Es liegen noch keine Erkenntnisse dazu vor, wie ein Tf entscheidet, welche Reaktion auf ein wahrgenommenes Objekt die richtige ist. Auch in den Regelwerken finden sich dazu keine Informationen. Weitere Kenntnisse über die Grundlagen der Entscheidungen von Tf bei der Objektwahrnehmung wären notwendig, um diese Entscheidungsprozesse genauer zu untersuchen und auch die Zeit zu bemessen, die für diese Entscheidungen notwendig ist. Ein halbstrukturiertes Interview zeigte, dass die Wahrnehmung des Risikos und das Verhalten von Tf von verschiedenen Faktoren beeinflusst werden können (Abschnitt 8.5). Zu diesem Zweck wäre es z. B.

möglich, vertiefende Interviews mit Tf über deren Risikowahrnehmung und Entscheidungsfindungsprozesse zu führen.

# 13 Abkürzungsverzeichnis

<b>Abkürzung</b>	<b>Bedeutung</b>
<b>AFT</b>	Accelerated Failure Time
<b>AP</b>	Arbeitspaket
<b>ATO</b>	Automatic Train Operation
<b>BBI</b>	Fachgebiet Bahnbetrieb und Infrastruktur der Technischen Universität Berlin
<b>BEU</b>	Bundesstelle für Eisenbahnunfalluntersuchung
<b>BÜ</b>	Bahnübergang
<b>DB</b>	Deutsche Bahn
<b>DIN</b>	Deutsches Institut für Normung
<b>DLR</b>	Deutsches Zentrum für Luft- und Raumfahrt e. V.
<b>DZSF</b>	Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt
<b>ECDF</b>	Empirical Cumulative Distribution Function (Empirische kumulative Verteilungsfunktion)
<b>ETCS</b>	European Train Control System
<b>FaS</b>	Fahren auf Sicht
<b>Fdl</b>	Fahrdienstleiter
<b>GoA</b>	Grade of Automation
<b>GoA3+</b>	GoA3 und GoA4
<b>HRA</b>	Human Reliability Assessment
<b>La</b>	Langsamfahrtstelle
<b>LiDAR</b>	Light detection and ranging
<b>LRT</b>	Likelihood-Ratio-Test
<b>LZB</b>	Linienförmige Zugbeeinflussung
<b>PSF</b>	Performance Shaping Factors
<b>PZB</b>	Punktförmige Zugbeeinflussung
<b>RADAR</b>	Radio detection and ranging
<b>Ril</b>	Richtlinie
<b>RT</b>	Reaktionszeit
<b>SiFa</b>	Sicherheitsfahrerschaltung
<b>Tf</b>	Triebfahrzeugführer (m/w/d)
<b>UFOV</b>	Useful Field of View
<b>VDE</b>	Verband der Elektrotechnik, Elektronik und Informationstechnik

# 14 Abbildungsverzeichnis

Abbildung 1: Aktualisierte Arbeitspakete und Zeitplan des ATO-Sense Projekts.....	12
Abbildung 2: Vorgehensweise .....	13
Abbildung 3: Einflussfaktoren auf die Leistungsfähigkeit der visuellen Wahrnehmung .....	20
Abbildung 4: Kontrastvarianten der Würfel der BBI-Simulatorstudie .....	38
Abbildung 5: Größen- und Kontrastvarianten der Würfel der DLR-Simulatorstudie.....	39
Abbildung 6: Die Simulationsumgebung der BBI.....	41
Abbildung 7: Die RailSET Simulationsumgebung.....	41
Abbildung 8: Anzahl der Datenpunkte (y-Achse) für jede Bedingung als Objektgröße. Geschwindigkeit. Objektfarbe. Zugsicherungssystem (x-Achse).....	48
Abbildung 9: Histogramm der Reaktionszeitdaten (links) und das Cullen-Frey-Diagramm (rechts).	50
Abbildung 10: Residuen des rohen Reaktionszeitmodells (links) und des log-transformierten Reaktionszeitmodells (rechts) für die DLR-Studie verglichen.....	50
Abbildung 11: Maße der zentralen Tendenz für einige Stimulus-Eigenschaftsbedingungen. Geometrische Mittelwerte sind in magenta, arithmetische Mittel in blau und Medianwerte in gelb.....	54
Abbildung 12: Interaktion zwischen Geschwindigkeit und Größe für kleine Stimuli (geometrische Mittelwerte als graue Punkte).....	56
Abbildung 13: Reaktionszeiten unter PZB 40 km/h (links) und unter FaS-Bedingungen (rechts) .....	57
Abbildung 14: Reaktionszeit bei verschiedenen Zugsicherungssystemen (links), verschiedenen Geschwindigkeitsklassen (rechts). Die Whisker stellen 3IQR dar, die grauen Punkte sind die geometrischen Mittelwerte. Reaktionszeit-Ausreißer wurden bei der Analyse nicht ausgeschlossen. ....	57
Abbildung 15: Vergleich der Modellvorhersage mit dem entsprechenden Indikator für die zentrale Tendenz unter verschiedenen Bedingungen (BBI-Studie). Log-transformierte Reaktionszeiten mit Stichprobe und vorhergesagten arithmetischen Mittelwerten (links) und rohen Reaktionszeiten mit Probe und vorhergesagten geometrischen Mittelwerten (rechts).....	59
Abbildung 16: Psychomotor-Vigilanz-Test (links) und Schläfrigkeitsbewertung (Mitte) über die Versuchsblöcke der BBI-Studie. Block 1 und Block 2 waren die PZB- und ETCS- Abschnitte in randomisierter Reihenfolge, während Block 3 immer das Fahren auf Sicht war. Bewertung der Arbeitsbelastung nach jeder Fahrt (rechts). Die Durchschnittswerte sind rot markiert. ....	60
Abbildung 17: Häufigkeit des Ereignisses: Große Objekte auf dem Lichtraumprofil (N=17).....	61
Abbildung 18: Häufigkeit des Ereignisses: Personen auf dem Lichtraumprofil (N=17).....	62
Abbildung 19: Maße der zentralen Tendenz für einige Stimulus-Eigenschaftsbedingungen (links) und Zugbetriebsparameterbedingungen (rechts) in der Simulatorstudie. Geometrische Mittelwerte sind gelb dargestellt, arithmetische Mittelgrün und Medianwerte rot.....	63
Abbildung 20: Histogramm der in der DLR-Simulatorstudie gemessenen Reaktionszeiten .....	64
Abbildung 21: Schläfrigkeitsbewertung über die Versuchsblöcke der DLR-Simulatorstudie. Die Durchschnittswerte sind grün markiert. Block 1 und Block 2 waren die PZB- und ETCS-	

Abschnitte in randomisierter Reihenfolge, während Block 3 immer das Fahren auf Sicht war. ....	67
Abbildung 22: Maße der zentralen Tendenz für einige Stimulus Eigenschaftsbedingungen (links) und Zugbetriebsparameterbedingungen (rechts) in der DLR-Simulatorstudie. Geometrische Mittelwerte sind gelb dargestellt, arithmetische Mittel grün und Medianwerte rot. ....	70
Abbildung 23: Diagramm der in der Online-Studie gemessenen Reaktionszeiten. ....	70
Abbildung 24: Vergleich der Modellvorhersage mit dem entsprechenden Indikator der zentralen Tendenz über die Bedingungen hinweg (Online-Studie). Logarithmisch transformierte Reaktionszeiten mit arithmetischen Mittelwerten der Stichprobe und der Vorhersage (links) und rohe Reaktionszeitdaten mit Probe und vorhergesagten geometrischen Mittelwerten (rechts).....	72
Abbildung 25: Histogramm der rohen und log-transformierten Rohreaktionszeiten (links) und empirische und theoretische kumulative Verteilungsfunktion der log-transformierten Reaktionszeitdaten (rechts) .....	76
Abbildung 26: Maße der zentralen Tendenz einige Szenarien .....	76
Abbildung 27: Reaktionszeiten für den gesamten Datensatz bei 40 km/h (oberes Feld), 100 km/h (mittleres Feld) und 160 km/h (unteres Feld). Y-Achse: Reaktionszeiten in Sekunden, X-Achse: Stimulusgröße in Bogenminuten. Die roten und gelben Punkte stellen die Beobachtungen für Stimuli mit niedrigem und hohem Kontrast dar.....	78
Abbildung 28: Einzelne Achsenabschnitte von Probanden, die innerhalb jedes Simulators geclustert wurden (Modell-1) (links), Einzelne Achsenabschnitte, wenn die Unterschiede in der Winkelgröße berücksichtigt wurden. (Modell-2) (rechts). ....	83
Abbildung 29: Histogramm der einzelnen Achsenabschnitte der Probanden (M2). ....	83
Abbildung 30: QQ-Plot zur Darstellung des Vergleichs der beobachteten Quantile mit den erwarteten Quantilen (links), Diagramm der Residuen und der angepassten Werte (rechts) .....	84
Abbildung 31: Cook's-Distanz-Plot zur Darstellung des Einflusses einzelner Beobachtungen auf das Regressionsmodell. Die rote Linie stellt den Schwellenwert $(4/n)$ dar. ....	85
Abbildung 32: Beobachtete Datenpunkte und die Modellvorhersagen basierend auf den individuellen Achsenabschnitten der Probanden für jeden Stimulus. X-Achse: individuelle Teilnehmenden-ID.....	87
Abbildung 33: Vergleich der Modellvorhersage mit dem entsprechenden Indikator der zentralen Tendenz über die Bedingungen hinweg. Y-Achse: Reaktionszeiten in Sekunden. X-Achse: Versuchsbedingungen als Kombination von Winkelgröße und Kontraststufen. ....	88
Abbildung 34: Reaktionszeitschätzung für die kontrastreichen Objekte im PZB-Szenario und die vom Benutzer eingegebene Bedingung (grüner Punkt). ....	90
Abbildung 35: Dichtefunktionen der verschiedenen Stufen der Variablen: Objektgröße (oben-links), Kontrast (oben-rechts), Geschwindigkeit (unten-links), Zubeinflussungssystem (unten-rechts) .....	91
Abbildung 36: Dichteverteilung der log-transformierten Reaktionszeiten in den Studien BBI und DLR (links) und Histogramm der Residuen des gesamten Datensatzes .....	92
Abbildung 37: Empirische kumulative Verteilungsfunktionen der verschiedenen Stufen der Variablen: Objektgröße (oben-links), Kontrast (oben-rechts), Geschwindigkeit (unten-links), Zugsicherungssystem (unten-rechts).....	93

Abbildung 38: Empirische kumulative Verteilungsfunktion (ECDF) der beobachteten Werte und vorhergesagten Werte für den gesamten Datensatz.....	95
Abbildung 39: Wahrscheinlichkeiten der Nichterkennung von Objekten für verschiedene Bedingungen. Die X-Achse ist die Zeit seit dem Erscheinen des Objekts. ....	97
Abbildung 40: Wahrscheinlichkeiten, die Objekte bei 40 km/h nicht zu erkennen.....	98

# 15 Tabellenverzeichnis

Tabelle 1: Standardaufgaben .....	15
Tabelle 2: Identifizierte Einflussfaktoren für die menschlichen Sinne .....	30
Tabelle 3: Standardaufgaben und deren jeweilige Automatisierbarkeit und Umsetzbarkeit im Simulator .....	33
Tabelle 4: Eigenschaften des wahrzunehmenden Reizes und Faktoren, die die Wahrnehmung beeinflussen .....	34
Tabelle 5: Unabhängige Variablen .....	35
Tabelle 6: Berechnung des Kontrastverhältnisses von zwei verschiedenen Bedingungen (BBI) .....	38
Tabelle 7: Beleuchtungsstärken der Würfel und Umgebung der DLR-Simulatorstudie .....	38
Tabelle 8: Experimentalbedingungen der drei Studien .....	44
Tabelle 9: Stichproben in den experimentellen Studien .....	47
Tabelle 10: Kontraste und Dummy-Kodierung für alle Faktoren der Simulatorstudien .....	52
Tabelle 11: Kontraste und Kodierung von Dummy-Variablen für alle Faktoren .....	53
Tabelle 12: Geometrische Mittelwerte und GEOMETRISCHE standardabweichungen (in Klammern gesetzt) der Reaktionszeit der verschiedenen Variablenkombinationen in der BBI-Simulatorstudie .....	55
Tabelle 13: Regressionskoeffizienten und t-Werte für die lineare Regression der Simulatorstudie ..	55
Tabelle 14: Mittelwerte und Standardabweichungen der Reaktionszeit der verschiedenen Variablenkombinationen in der DLR-Simulatorstudie .....	64
Tabelle 15: Regressionskoeffizienten und t-Werte für die lineare Regression der DLR-Simulatorstudie. Multiple R-squared: 0.4559, Adjusted R-squared: 0.4481 .....	65
Tabelle 16: Mittelwerte und Standardabweichungen der verschiedenen Variablenkombinationen in der Online-Studie .....	69
Tabelle 17: Regressionskoeffizienten und t-Werte für die lineare Regression der Onlinestudie .....	71
Tabelle 18: Scheinbare Größe der Stimuli auf BBI- und DLR-Simulatoren .....	75
Tabelle 19: Geometrische Mittelwerte der Reaktionszeit der verschiedenen Variablenkombinationen .....	77
Tabelle 20: Kodierung von Dummy-Variablen für alle Faktoren .....	80
Tabelle 21: Schätzungen des Modells mit gemischten Effekten (M1) und Modellparameter .....	80
Tabelle 22: Dummy-Variable für die Winkelgröße und die paarweisen Vergleiche durch Kontrastkodierung .....	82
Tabelle 23: Schätzungen des Modells mit gemischten Effekten (M2) .....	82
Tabelle 24: Vergleich der Zufallseffekte und Modellparameter der Modelle M1 und M2 .....	84
Tabelle 25: Standard-Fehler (mit und ohne Random-Effekte) .....	86
Tabelle 26: Modellparameter der Trainings- und Validierungsdatensätze .....	86
Tabelle 27: Modellparameter der Regressionsanalyse M9 .....	89
Tabelle 28: Kolmogorow-Smirnow-Test-Parameter für verglichene Paare .....	94

Tabelle 29: Schätzungen des Weibull-AFT-Modells ..... 96

## 16 Quellenverzeichnis

- ANDERSSON, M., G. BJÖRKLUND und M. HARALDSSON, 2016. Marginal railway track renewal costs: A survival data approach [online]. *Transportation Research Part A: Policy and Practice*, **87**, 68-77. ISSN 0965-8564. Verfügbar unter: doi:10.1016/j.tra.2016.02.009
- BAAAYEN, R.H. und P. MILIN, 2010. Analyzing Reaction Times [online]. *International Journal of Psychological Research*, **3**(2), 12-28. ISSN 2011-2084. Verfügbar unter: <https://dialnet.unirioja.es/servlet/articulo?codigo=3405162>
- BARTMANN, A., D. REIFFENRATH, A.M. JACOBS, H. LEDER, M. WALKOWIAK und A. SZYMKOWIAK, 1993. *Sichtabstand bei Fahrten in der Dunkelheit*: Verlag für neue Wissenschaft. Berichte der Bundesanstalt für Straßenwesen, Reihe Mensch und Sicherheit. M4.
- BATES, D., M. MÄCHLER, B. BOLKER und S. WALKER, 2015. Fitting Linear Mixed-Effects Models Using lme4 [online]. *Journal of Statistical Software*, **67**(1). Journal of Statistical Software. Verfügbar unter: doi:10.18637/jss.v067.i01
- BECKER-CARUS, C. und M. WENDT, 2016. *Allgemeine Psychologie. Eine Einführung* [online]. Berlin: Springer. ISBN 978-3-662-53006-1. Verfügbar unter: <http://www.springer.com/>
- BECKER-CARUS, C. und M. WENDT, 2017. Allgemeine Psychologie [online]. Verfügbar unter: doi:10.1007/978-3-662-53006-1
- BEU, o. J. *Jahresberichte 2014-2020* [online]. *Berichte über die Arbeit der Eisenbahn-Unfalluntersuchung des Bundes*. Verfügbar unter: <https://www.eisenbahn-unfalluntersuchung.de/EUB/DE/Publikationen.html>
- BOFF, K. und J. LINCOLN, 1988. *Engineering data compendium. Human perception and performance*. Wright-Patterson Air Force Base, OH: Armstrong Medical Research Laboratory. 19880019248.
- BONNET, C., J. GURLEKIAN und P. HARRIS, 1992. Reaction time and visual area: Searching for the determinants [online]. *Bulletin of the Psychonomic Society*, **30**(5), 396-398. Bulletin of the Psychonomic Society. Verfügbar unter: doi:10.3758/BF03334099
- BORING, R., 2010. *How Many Performance Shaping Factors are Necessary for Human Reliability Analysis* [online]. University of North Texas Libraries, UNT Digital Library. Idaho Falls, Idaho [Zugriff am: 1. September 2021]. Verfügbar unter: <https://digital.library.unt.edu/ark:/67531/metadc835988/>
- BORING, R.L., 2012. *Fifty Years of THERP and Human Reliability Analysis* [online]. Idaho National Lab. (INL), Idaho Falls, ID (United States). INL/CON-12-25623. Verfügbar unter: <https://www.osti.gov/biblio/1082366>
- BOROWSKY, A., W.J. HORREY, Y. LIANG, A. GARABET, L. SIMMONS und D.L. FISHER, 2016. The effects of brief visual interruption tasks on drivers' ability to resume their visual search for a precued hazard [online]. *Accident; analysis and prevention*, **93**, 207-216. Accident; analysis and prevention. Verfügbar unter: doi:10.1016/j.aap.2016.04.028

- BRABAND, J., B. EVERS, M. KINAS, L. LINDNER, D. MIHAILESCU-STOICA, F. REXIN, F. ADEBAHR, B. MILIUS und H. SCHÄBE, 2023. *Risikoakzeptanzkriterien für das automatisierte Fahren auf der Schiene* [online]. Deutsches Zentrum für Schienenverkehrsforschung. Berichte des Deutschen Zentrums für Schienenverkehrsforschung. Bericht 40 (2023), Projektnummer 2020-19-S-1202. Verfügbar unter: [https://www.dzsf.bund.de/SharedDocs/Fachmitteilungen/DZSF/2023/08\\_2023\\_Forschungsbericht\\_40\\_2023.html](https://www.dzsf.bund.de/SharedDocs/Fachmitteilungen/DZSF/2023/08_2023_Forschungsbericht_40_2023.html)
- BRANDENBURGER, N., A. NAUMANN und M. JIPP, 2019. Task-induced fatigue when implementing high grades of railway automation [online]. *Cognition, Technology & Work*. ISSN 1435-5558. Verfügbar unter: doi:10.1007/s10111-019-00613-z
- BRANDENBURGER, N., B. THOMAS-FRIEDRICH, A. NAUMANN und J. GRIPPENKOVEN, Hg., 2018. *Automation in railway operations: effects on signaller and train driver workload*. Proceedings of the 3rd German Workshop on Rail Human Factors (pp. 51-60).
- BRANDENBURGER, N., H.-J. HÖRMANN, D. STELLING und A. NAUMANN, 2017. Tasks, skills, and competencies of future high-speed train drivers [online]. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **231**(10), 1115-1122. ISSN 0954-4097. Verfügbar unter: doi:10.1177/0954409716676509
- BRANDENBURGER, N.N., M. WITTKOWSKI und A. NAUMANN, 2017. Countering Train Driver Fatigue in Automatic Train Operation [online]. In: D. GOLIGHTLY, A. FOWLER, B. RYAN, S. KURUP und A. MILLS, Hg. *Sixth International Human Factors Rail Conference*: Rail Safety and Standards Board, S. 57-67. Verfügbar unter: <https://elib.dlr.de/113637/>
- CAO, Y. und J.-H. WANG, 2004. A Driving Simulation Based Study on the Effects of Road Marking Luminance Contrast on Driving Safety [online]. In: *Institute of Transportation Engineers*. Burlington, Vermont. Verfügbar unter: <https://www.semanticscholar.org/paper/Title%3A-A-Driving-Simulation-Based-Study-on-the-of-Cao/32d053e8a7ca4f8d28bdabf15e8509c92c4b7350>
- CAVALLO, V. und A.S. COHEN, 2001. Perception. In: P.E. BARJONET, Hg. *Traffic psychology today*: Kluwer Academic Publisher, S. 63-89.
- CHERRY, C.E., 1953. Some Experiments on the Recognition of Speech, with One and with Two Wars. *The Journal of the Acoustical Society of America*, **25**(5), 975-979. The Journal of the Acoustical Society of America.
- CHOUDHARY, P. und N.R. VELAGA, 2017. Modelling driver distraction effects due to mobile phone use on reaction time [online]. *Transportation Research Part C: Emerging Technologies*, **77**, 351-365. ISSN 0968-090X. Verfügbar unter: doi:10.1016/j.trc.2017.02.007
- CLARK, T.G., M.J. BRADBURN, S.B. LOVE und D.G. ALTMAN, 2003. Survival analysis part I: basic concepts and first analyses [online]. *British Journal of Cancer*, **89**(2), 232-238. ISSN 1532-1827. Verfügbar unter: doi:10.1038/sj.bjc.6601118
- COHEN, A.S., 1987. Blickverhalten des Fahrzeuglenkers als Komponente des Verkehrssicherheitsverhaltens. In: H. HÄCKER, Hg. *Fortschritte der Verkehrspsychologie*: TÜV Rheinland, S. 112-125.
- COHEN, A.S., 1993. Visuelle Informationsaufnahme des motorisierten Verkehrsteilnehmers. *Verkehrsunfall und Fahrzeugtechnik*, **31**(7/8), 209-214. Verkehrsunfall und Fahrzeugtechnik.

- COHEN, A.S., 2008. Wahrnehmung als Grundlage der Verkehrsorientierung bei nachlassender Sensorik während der Alterung. In: B. SCHLAG, Hg. *Leistungsfähigkeit und Mobilität im Alter*: Verlag TÜV Rheinland., S. 65-84.
- COHEN, J. und P. COHEN, 2013. *Applied multiple regression-correlation analysis for the behavioral sciences*. Second edition: Routledge. ISBN 9780203774441.
- COOK, R.D., 1977. Detection of Influential Observation in Linear Regression [online]. *Technometrics*, **19**(1), 15. ISSN 00401706. Verfügbar unter: doi:10.2307/1268249
- COOK, R.D., 2011. Cook's Distance. In: M. LOVRIC, Hg. *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 301-302. ISBN 978-3-642-04898-2.
- CRUNDALL, D.E. und G. UNDERWOOD, 1998. Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, **41**(4), 448-458. Ergonomics.
- CZAMOLEWSKI, M., 1996. An Empirical Validation for the Natural Log Transformation of Reaction Time [online]. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **40**(24), 1287. ISSN 2169-5067. Verfügbar unter: doi:10.1177/154193129604002471
- DACOTA, 2012. *Fatigue, Deliverable 4.8h*. EC FP7 Project DaCoTA.
- DB Netz AG. Richtlinie 408.21-27:2014, *Neuherausgabe Richtlinie 408 - Fahrdienstvorschrift*: DB Netz AG.
- DESMOND, P.A., P.A. HANCOCK und J.L. MONETTE, 1998. Fatigue and Automation-Induced Impairments in Simulated Driving Performance [online]. *Transportation Research Record: Journal of the Transportation Research Board*, **1628**(1), 8-14. ISSN 0361-1981. Verfügbar unter: doi:10.3141/1628-02
- Deutsches Institut für Normung e. V. DIN VDE V 0831-103:2020, *Elektrische Bahn-Signalanlagen - Teil 103: Ermittlung von Sicherheitsanforderungen an technische Funktionen in der Eisenbahnsignaltechnik*.
- DIE EUROPÄISCHE KOMMISSION, 2013. *Durchführungsverordnung (EU) Nr. 402/2013 der Kommission vom 30. April 2013 über die gemeinsame Sicherheitsmethode für die Evaluierung und Bewertung von Risiken und zur Aufhebung der Verordnung (EG) Nr. 352/2009*. 402/2013. L 121.
- DIGGLE, P., P. HEAGERTY, K.-Y. LIANG und S.L. ZEGER, 2013. *Analysis of longitudinal data* [online]. 2nd ed. Oxford: Oxford University Press. Oxford statistical science series. no. 25. ISBN 0191664324. Verfügbar unter: <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=598873>
- DIN. DIN EN 16186-1:2019, *Bahnanwendungen – Führerraum*: Deutsches Institut für Normung.
- DINGES, D.F. und J.W. POWELL, 1985. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations [online]. *Behavior Research Methods, Instruments, & Computers*, **17**(6), 652-655. ISSN 1532-5970. Verfügbar unter: doi:10.3758/BF03200977
- DIRECTIVE (EU) 2016/798 of the European Parliament and of the Council of 11 May 2016 on railway safety. In: *Official Journal of the European Union*.

- DORRIAN, J., G.D. ROACH, A. FLETCHER und D. DAWSON, 2007. Simulated train driving: fatigue, self-awareness and cognitive disengagement [online]. *Applied ergonomics*, **38**(2), 155-166. ISSN 0003-6870. Verfügbar unter: doi:10.1016/j.apergo.2006.03.006
- ECKERT, M., 1993. *Lichttechnik und optische Wahrnehmungssicherheit im Straßenverkehr*: Technik.
- ELLINGHAUS, D., B. SCHLAG und J. STEINBRECHER, 1990. *Leistungsfähigkeit und Fahrverhalten älterer Kraftfahrer*: Wirtschaftsverlag NW, Verlag für neue Wissenschaft. Berichte der Bundesanstalt für Straßenwesen, Unfall- und Sicherheitsforschung. M80.
- FARAWAY, J.J., 2016. *Extending the linear model with R. Generalized linear, mixed effects and non-parametric regression models*. Second edition. Boca Raton, Florida: CRC P. Texts in statistical science. ISBN 9781315382722.
- GELMAN, A. und J. HILL, 2009. *Data analysis using regression and multilevel/hierarchical models*. 10. print. Cambridge: Cambridge Univ. Press. Analytical methods for social research. ISBN 9780521867061.
- GIBSON, J.J., 1962. Observations on active touch. *Psychological Review*, **69**(6), 477-491. Psychological Review.
- GOLDSTEIN, E.B., 1996. *Sensation and perception*: Brooks/Cole Publishing Company.
- GOLDSTEIN, E.B., 2008. *Wahrnehmungspsychologie*: Spektrum.
- GRAMBERG-DANIELSEN, B., 1967. *Sehen und Verkehr*: Springer.
- GRÜNDL, M., 2005. *Fehler und Fehlverhalten als Ursache von Verkehrsunfällen und Konsequenzen für das Unfallvermeidungspotenzial und die Gestaltung von Fahrerassistenzsystemen*. Dissertation. Regensburg.
- GUO, B., Y. MAO, A. HEDGE und W. FANG, 2015. Effects of apparent image velocity and complexity on the dynamic visual field using a high-speed train driving simulator [online]. *International Journal of Industrial Ergonomics*, **48**, 99-109. ISSN 01698141. Verfügbar unter: doi:10.1016/j.ergon.2015.04.005
- HAEGERSTROM-PORTNOY, G., M.E. SCHNECK und J.A. BRABYN, 1999. Seeing into Old Age. Vision Function Beyond Acuity. *Optometry and Vision Science*, **76**(3). Optometry and Vision Science.
- HAKKINEN, S., 1963. *Estimation of distance and velocity judgments in traffic situations*. Helsinki, Finland: Institute of Occupational Health.
- HAMMERL, M. und F. VANDERHAEGEN, Hg., 2009. *Human factors in the railway system safety analysis process*.
- HAQUE, M.M. und S. WASHINGTON, 2014. A parametric duration model of the reaction times of drivers distracted by mobile phone conversations [online]. *Accident; analysis and prevention*, **62**, 42-53. Accident; analysis and prevention. Verfügbar unter: doi:10.1016/j.aap.2013.09.010
- HEINZ, G., 1982. Einige Ursachen für das erhöhte Unfallrisiko bei Dunkelheit. *Straßenverkehrstechnik*, **6**, 179-183. Straßenverkehrstechnik.

- HELY, M., T. SHARDLOW, B.R. BUTT, R. FRISWELL, A. MCINTOSH und AND ANN WILLIAMSON., 2015. Effects of automatic train protection on human factors and driving behaviour. *Proceedings 19th Triennial Congress of the IEA, Melbourne 9-14 August 2015*. Proceedings 19th Triennial Congress of the IEA, Melbourne 9-14 August 2015.
- HENTSCHEL, H.-J., 2002. *Licht und Beleuchtung. Grundlagen und Anwendungen der Lichttechnik*. 5. Auflage. Heidelberg: Hüthig.
- HÖH, A.E., T. ACH, R. AMBERGER und S. DITHMAR, 2008. Lichtexposition bei vitreoretinaler Chirurgie. I. Grundlagen [online]. *Der Ophthalmologe : Zeitschrift der Deutschen Ophthalmologischen Gesellschaft*, **105**(10), 898-900, 902-4. ISSN 0941-293X. Verfügbar unter: doi:10.1007/s00347-008-1794-z
- IEC. DIN 62267:2010, *Bahnanwendungen - Automatischer städtischer schienengebundener Personennahverkehr (AUGT) Sicherheitsanforderungen*.
- JACKSON, C.H., 2016. flexsurv: A Platform for Parametric Survival Modeling in R [online]. *Journal of statistical software*, **70**. ISSN 1548-7660. Verfügbar unter: doi:10.18637/jss.v070.i08
- JOHNE, M. und M. BUSSE, 2016. RailSiTe® (Rail Simulation and Testing) [online]. *Journal of large-scale research facilities JLSRF*, **2**. Journal of large-scale research facilities JLSRF. Verfügbar unter: doi:10.17815/jlsrf-2-144
- JONIDES, J., 1981. Voluntary versus automatic control over the mind's eye's movement. In: LONG, J., BADDELEY, A., Hg. *Attention and Performance*. Hillsdale, NJ: Erlbaum, S. 187-203.
- KLEIN, J.P. und M.L. MOESCHBERGER, 2003. *Survival analysis: techniques for censored and truncated data*: Springer.
- KLEINBAUM, D.G. und M. KLEIN, 2012. *Survival analysis. A self-learning text* [online]. 3rd ed. New York, NY: Springer. Statistics for biology and health. ISBN 1441966463. Verfügbar unter: <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=2546410>
- KLINKE, R., 1995. Der Gleichgewichtssinn. In: R.F. SCHMIDT und G. THEWS, Hg. *Physiologie des Menschen*. Sechszwanzigste, vollständig überarbeitete Auflage. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 251-257. ISBN 978-3-662-09334-4.
- KRAMER, A.F., 1990. *Physiological Metrics of Mental Workload. A Review of Recent Progress*. San Diego. NPRDC-TN-90-23.
- KURATORIUM FÜR VERKEHRSSICHERHEIT, 2006. *Ermüdung als Unfallursache im Straßenverkehr* [online]. Verfügbar unter: <http://www.kfv.at/index.php?id=602>
- KYRIAKIDIS, M., A. MAJUMDAR, G. GROTE und W.Y. OCHIENG, 2012. Development and Assessment of Taxonomy for Performance-Shaping Factors for Railway Operations [online]. *Transportation Research Record: Journal of the Transportation Research Board*, **2289**(1), 145-153. ISSN 0361-1981. Verfügbar unter: doi:10.3141/2289-19
- LACHENMAYR, B., 1995. *Sehen und gesehen werden. Sicher unterwegs im Straßenverkehr.*: Shaker. Berichte aus der Medizin.

- LACHENMAYR, B., 2003. Tagfahrlicht [online]. Pro und contra. *Ophthalmologe*, **100**(10), 836-842. Ophthalmologe. Verfügbar unter: doi:10.1007/s00347-002-0755-1
- LACHENMAYR, B., 2006. Gesichtsfeld und Verkehr. Wie funktioniert das periphere Sehen? [online]. *Ophthalmologe*, **103**(5), 373-381. Ophthalmologe. Verfügbar unter: doi:10.1007/s00347-005-1268-5
- LAMBERT, A., E. SPENCER und N. MOHINDRA, 1987. Automaticity and the capture of attention by a peripheral display change. *Current Psychology*, **6**(2), 136-147. Current Psychology.
- LAND, M.F. und J.M. HORWOOD, 1995. Which Parts of the Road Guide Steering? *Nature*, **377**, 339-340. Nature.
- LEE, Y. und J.A. NELDER, 2004. Conditional and Marginal Models: Another View [online]. *Statistical Science*, **19**(2). ISSN 0883-4237. Verfügbar unter: doi:10.1214/088342304000000305
- LEEUW, J.R. de, 2015. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser [online]. *Behavior Research Methods*, **47**(1), 1-12. ISSN 1554-3528. Verfügbar unter: doi:10.3758/s13428-014-0458-y
- LEINHOS, D., H. FLATT und S. WITTE, 2022. *Sensorik als technische Voraussetzung für ATO-Funktionen. Berichte des Deutschen Zentrums für Schienenverkehrsforschung, Bericht 31 (2022)*. Deutsche Zentrum für Schienenverkehrsforschung. Dresden. ISSN 2629-7973.
- LINDNER, T., B. MILIUS, M. ARENIUS, D. SCHWENCKE, J. GRIPPENKOVEN und O. STRÄTER, 2014. Betrachtungen zur Zuverlässigkeit des Triebfahrzeugführers. Erfassung sicherheitsbeeinflussender Faktoren und ihrer Bedeutung auf Basis von Ereignisdaten. *EI - Der Eisenbahningenieur*, ((1/2014)), 10-16. EI - Der Eisenbahningenieur.
- LIU, Y., R. ALSALEH und T. SAYED, 2021. Modeling the influence of mobile phone use distraction on pedestrian reaction times to green signals: A multilevel mixed-effects parametric survival model [online]. *Transportation Research Part F: Traffic Psychology and Behaviour*, **81**, 115-129. ISSN 1369-8478. Verfügbar unter: doi:10.1016/j.trf.2021.05.020
- MARINAKOS, H., T.B. SHERIDAN und J. MULTER, 2005. *Effects of Supervisory Train Control Technology on Operator Attention*. U.S. Department of Transportation. Human Factors in Railroad Operations. DOT/FRA/ORD-04/10.
- MCGURK, H. und J. MACDONALD, 1976. Hearing lips and seeing voices. *Nature*, **264**, 746-748. Nature.
- MCLEOD, R.W., G.H. WALKER und N. MORAY, 2005. Analysing and modelling train driver performance [online]. *Applied ergonomics*, **36**(6), 671-680. ISSN 0003-6870. Verfügbar unter: doi:10.1016/j.apergo.2005.05.006
- MIURA, T., 1992. Visual search in intersections and underlying mechanism. *IATSS Research*, **16**(1), 42-49. IATSS Research.
- MOKRZYCKI, W.S. und M. TATOL, 2011. Colour Difference  $\Delta E$  - a Survey. *MG&V*, **20**(4), 383-411. ISSN 1230-0535.

- MOURANT, R.R. und T.H. ROCKWELL, 1972. Strategies of Visual Search by Novice and Experienced Drivers. *Human Factors*, **14**(4), 325-335. Human Factors.
- MROWINSKI, D., G. SCHOLZ und T. STEFFENS, 2017. *Audiometrie*. Stuttgart: Georg Thieme Verlag. ISBN 9783132401075.
- NAGHIYEV, A., S. SHARPLES, M. CAREY, A. COPLESTONE und B. RYAN, 2014. ERTMS train driving-incab vs. outside. In: S. SHARPLES und S. SHORROCK, Hg. *Contemporary Ergonomics and Human Factors 2014*: Taylor & Francis, S. 343-350. ISBN 978-1-138-02635-3.
- OECD / ECMT, 2006. *Speed Management*.
- PAVLOU, M., G. AMBLER, S. SEAMAN und R.Z. OMAR, 2015. A note on obtaining correct marginal predictions from a random intercepts model for binary outcomes [online]. *BMC Medical Research Methodology*, **15**(1), 59. ISSN 1471-2288. Verfügbar unter: doi:10.1186/s12874-015-0046-6
- PAWAR, N.M. und N.R. VELAGA, 2020. Modelling the influence of time pressure on reaction time of drivers [online]. *Transportation Research Part F: Traffic Psychology and Behaviour*, **72**, 1-22. ISSN 1369-8478. Verfügbar unter: doi:10.1016/j.trf.2020.04.017
- PIÉRON, H., 1914. II. Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices [online]. *L'Année psychologique*, **20**(1), 17-96. ISSN 0003-5033. Verfügbar unter: doi:10.3406/psy.1913.4294
- PINS, D. und C. BONNET, 1996. On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time [online]. *Perception & Psychophysics*, **58**(3), 390-400. Perception & Psychophysics. Verfügbar unter: doi:10.3758/BF03206815
- PLEWAN, T. und G. RINKENAUER, 2017. Simple reaction time and size-distance integration in virtual 3D space [online]. *Psychological research*, **81**(3), 653-663. Psychological research. Verfügbar unter: doi:10.1007/s00426-016-0769-y
- PURVES, D. und S.M. WILLIAMS, 2004. *Neuroscience*. 3. Auflage. Sinauer: Sunderland and Mass.
- RATCLIFF, R., 1993. Methods for dealing with reaction time outliers [online]. *Psychological bulletin*, **114**(3), 510-532. ISSN 0033-2909. Verfügbar unter: doi:10.1037/0033-2909.114.3.510
- REMYNGTON, R.W., J.C. JOHNSTON und S. YANTIS, 1992. Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics*, **51**(3), 279-290. Perception & Psychophysics.
- RICHTER, P. und W. HACKER, 1998. *Belastung und Beanspruchung. Stress und Monotonie in der Arbeitstätigkeit*: Asanger.
- RICHTER, S. und B. SCHLAG, 2000. *Psychologische Untersuchungen zu Nebelunfällen*: Pabst Science Publishers.
- ROGÉ, J., T. PÉBAYLE, E. LAMBILLIOTTE, F. SPITZENSTETTER, D. GISELBRECHT und A. MUZET, 2004. Influence of age, speed and duration of monotonous driving task in traffic on the driver's useful visual field [online]. *Vision Research*, **44**(23), 2737-2744. ISSN 0042-6989. Verfügbar unter: doi:10.1016/j.visres.2004.05.026

- RSSB, 2005. *Rail-Specific Human Reliability Assessment Technique for Driving Tasks (T270)*. Rail Safety and Standards Board. T270.
- RSSB, 2012. *Railway Action Reliability Assessment User Manual – A Technique for the Quantification of Human Error in the Rail Industry*. Rail Safety and Standards Board.
- SANDERS, A.F., 1998. *Elements of human performance. Reaction processes and attention in human skill*. Mahwah, N.J.: Lawrence Erlbaum Associates. ISBN 9780203774250.
- SCHAALJE, G.B., J.B. MCBRIDE und G.W. FELLINGHAM, 2002. Adequacy of approximations to distributions of test statistics in complex mixed linear models [online]. *Journal of Agricultural, Biological, and Environmental Statistics*, **7**(4), 512-524. ISSN 1085-7117. Verfügbar unter: doi:10.1198/108571102726
- SCHLAG, B. und K.J. BECKMANN, Hg., 2013. *Mobilität und demografische Entwicklung*. Köln: TÜV Media. Mobilität und Alter. 7.
- SCHLAG, B. und R. HEGER, 2002. *Empfehlungen zur Berücksichtigung physiologischer und psychologischer Fähigkeiten und Grenzen der Kraftfahrer bei der Straßenplanung in Brandenburg. 1. Stufe: Systematisierung und Evaluation vorhandener Erkenntnisse*. Technische Universität Dresden. Unveröffentlichter Bericht.
- SCHLAG, B., Hg., 2008. *Leistungsfähigkeit und Mobilität im Alter*: Verlag TÜV Rheinland.
- SCHLAG, B., I. PETERMANN, C. SCHULZE und G. WELLER, Hg., 2009. *Mehr Licht - mehr Sicht - mehr Sicherheit? Zur Wirkung verbesserter Licht- und Sichtbedingungen auf das Fahrerverhalten*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden. Verkehrspsychologie. ISBN 978-3-531-91397-1.
- SCHMIDT-CLAUSEN, H.-J. und A. FREIDING, 2004. *Sehvermögen von Kraftfahrern und Lichtbedingungen im nächtlichen Straßenverkehr*. Bremerhaven: Verlag für neue Wissenschaft. Berichte der Bundesanstalt für Straßenwesen, Mensch und Sicherheit. M158.
- SEARLE, S.R., C.E. MCCULLOCH und G. CASELLA, 1992. *Variance components*. New York: Wiley. Wiley series in probability and mathematical statistics. Applied probability and statistics. ISBN 9780471621621.
- SHAHID, A., K. WILKINSON, S. MARCU und C.M. SHAPIRO, 2012. Karolinska Sleepiness Scale (KSS). In: A. SHAHID, K. WILKINSON, S. MARCU und C.M. SHAPIRO, Hg. *STOP, THAT and One Hundred Other Sleep Scales*. s.l.: Springer-Verlag, S. 209-210. ISBN 978-1-4419-9892-7.
- SPRENGER, A., 2001. Müdigkeit und Müdigkeitsmessung bei Lkw-Fahrerinnen und - Fahrern. Ein Überblick zum Stand der Forschung. *ILS-Schriften*, **176**, 36-40. ILS-Schriften.
- STANGL, W., 2021. *Stichwort: 'Webersches Gesetz – Online Lexikon für Psychologie und Pädagogik'*. *Online Lexikon für Psychologie und Pädagogik*. [online] [Zugriff am: 7. April 2021]. Verfügbar unter: <https://lexikon.stangl.eu/4639/webersches-geset>
- STEIN, J. und A. NAUMANN, 2016. Monotony, Fatigue and Microsleeps - Train driver' daily routine: a Simulator study [online]. *Rail Human Factors. Proceedings of the 2nd German Workshop on Rail Human Factors, Milius, B; Naumann, A (Eds.)*, 96-102. Rail Human Factors. Proceedings of

- the 2nd German Workshop on Rail Human Factors, Milius, B; Naumann, A (Eds.). Verfügbar unter: <https://elib.dlr.de/103841/>
- SUMMERFIELD, C. und T. EGNER, 2009. Expectation (and attention) in visual cognition [online]. *Trends in cognitive sciences*, **13**(9), 403-409. ISSN 1364-6613. Verfügbar unter: doi:10.1016/j.tics.2009.06.003
- SUZUKI, D., K. YAMAUCHI und S. MATSUURA, 2019. Effective Visual Behavior of Railway Drivers for Recognition of Extraordinary Events [online]. *Quarterly Report of RTRI*, **60**(4), 286-291. ISSN 0033-9008. Verfügbar unter: doi:10.2219/rtriqr.60.4\_286
- SWAIN, A. und H.E. GUTTMAN, 1983. *Handbook of human reliability analysis with emphasis on nuclear power plant applications. Final report*. NUREG/CR-1278: US Nuclear Regulatory Commission. Washington, DC. doi:10.2172/5752058.
- TUKEY, J.W., 1977. *Exploratory data analysis*: Reading, MA. 2.
- UNDERWOOD, G., P. CHAPMAN, K. BOWDEN und D. CRUNDALL, 2002. Visual search while driving: skill and awareness during inspection of the scene [online]. *Transportation Research Part F: Traffic Psychology and Behaviour*, **5**(2), 87-97. ISSN 1369-8478. Verfügbar unter: doi:10.1016/S1369-8478(02)00008-6
- VDI 4006. *Blatt 1: Menschliche Zuverlässigkeit – Ergonomische Forderungen und Methoden der Bewertung*.
- WAGENMAKERS, E.-J. und S. BROWN, 2007. On the linear relation between the mean and the standard deviation of a response time distribution [online]. *Psychological Review*, **114**(3), 830-841. Psychological Review. Verfügbar unter: doi:10.1037/0033-295X.114.3.830
- WALKER, B.H., 2009. *Optical Engineering Fundamentals, Second Edition*: SPIE. ISBN 9780819478771.
- WASHINGTON, S., 2020. *Statistical and Econometric Methods for Transportation Data Analysis*. 3rd ed. Milton: CRC Press LLC. Chapman and Hall/CRC Interdisciplinary Statistics Ser. ISBN 9780429520754.
- WELHAM, S., B. CULLIS, B. GOGEL, A. GILMOUR und R. THOMPSON, 2004. Prediction in linear mixed models [online]. *Australian & New Zealand Journal of Statistics*, **46**(3), 325-347. ISSN 1369-1473. Verfügbar unter: doi:10.1111/j.1467-842X.2004.00334.x
- WELLER, G., B. SCHLAG, G. GATTI, R. JORNA und M. VAN DE LEUR, 2006. *Human Factors in Road Design. State of the art and empirical evidence*. Report RITUD-WP8-R1-V5-Human-Factors.
- WICKENS, C.D. und J.G. HOLLANDS, 1999. *Engineering Psychology and Human Performance*. 3. Auflage. Upper Saddle River, New Jersey: Prentice-Hall INC.
- WICKENS, C.D., 2002. Multiple resources and performance prediction [online]. *Theoretical Issues in Ergonomics Science*, **3**(2), 159-177. ISSN 1463-922X. Verfügbar unter: doi:10.1080/14639220210123806

- WOLFE, B., J. DOBRES, R. ROSENHOLTZ und B. REIMER, 2017. More than the Useful Field: Considering peripheral vision in driving [online]. *Applied ergonomics*, **65**, 316-325. ISSN 0003-6870. Verfügbar unter: doi:10.1016/j.apergo.2017.07.009
- WÖRDENWEBER, B., 2004. *Mensch im Verkehr. Seminar 3b, Einführung in die Lichttechnik.*, 2004.
- X2RAIL-1, 2019. *Deliverable D4.3. ATO over ETCS – GoA3/4 Preliminary Specification* [online]. H2020-S2RJU-CFM-2015-01-1. Verfügbar unter: [https://projects.shift2rail.org/s2r\\_ip\\_TD\\_D.aspx?ip=2&td=b47388a9-b1f8-4ed8-9872-bb7708f7c08d](https://projects.shift2rail.org/s2r_ip_TD_D.aspx?ip=2&td=b47388a9-b1f8-4ed8-9872-bb7708f7c08d)
- YANTIS, S. und J. JONIDES, 1984. Abrupt Visual Onsets and Selective Attention. Evidence From Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, **10**(5), 601-621. *Journal of Experimental Psychology: Human Perception and Performance*.
- YEH, M. und J. MULTER, 2008. *Driver Behavior at Highway-Railroad Grade Crossings. A Literature Review Final Report from 1990-2006*. U.S. Department of Transportation. DOT/FRA/ORD-08/03.
- ZIEGEL, E.R. und L. OTT, 1994. An Introduction to Statistical Methods and Data Analysis [online]. *Technometrics*, **36**(3), 332. ISSN 00401706. Verfügbar unter: doi:10.2307/1269399

# 17 Anhänge

ANHANG 1: PERFORMANCE SHAPING FACTORS .....	121
ANHANG 2: PSF-EXPERTENSCHÄTZUNG (.XLSX DATEI) .....	123
ANHANG 3: R-TOOL (.R und .txt Dateien).....	123

ANHANG 1: PERFORMANCE SHAPING FACTORS

Kategorie	ID	PSF	Definition
Personal	PSF1	Alter	Alter des Fahrers
	PSF2	Geschlecht	Geschlecht des Fahrers
	PSF3	Gesundheit	Zustand des vollständigen körperlichen, geistigen und sozialen Wohlbefindens
	PSF4	Intelligenz/Bildungsniveau	Fähigkeit abstrakt und vernünftig zu denken und daraus zweckvolles Handeln abzuleiten; Bildungsstand
	PSF5	Anspannung	eine psychische oder physische Belastung, aus der Nervosität hervorgeht
Personal (beeinflusste Faktoren)	PSF6	Situationsbewusstsein	die Wahrnehmung der Elemente innerhalb eines Zeit- und Raumvolumens, das Verständnis ihrer Bedeutung und die Projektion ihres Status in naher Zukunft (Endsley, 1995).
	PSF7	Müdigkeit	ein geistiger oder körperlicher Zustand, der auf unzureichenden Schlaf, hohen oder längeren Arbeitsbedarf und einen zirkadianen Rhythmus zurückzuführen ist und die Fähigkeit der Arbeitnehmer beeinträchtigt, sicher zu arbeiten.
	PSF8	Ablenkung	unkritische sekundäre Aufgaben oder externe Einflüsse, die die Leistung der primären Aufgabe beeinflussen.
	PSF9	Risikobewusstsein	situationsabhängiges Bewusstsein der Handlungen und ihrer Folgen einschließlich Über- oder Unterschätzung von Risiken, Selbstzufriedenheit
	PSF10	Eignung	körperliche oder geistige Eignung für den Dienst
	PSF11	Stress	starke Beanspruchung, erschwert adäquate Durchführung der Aufgabe
	PSF12	Dichte Kopplung	hohe Interdependenz der Ereignisse oder Unteraufgaben (Kognitive Kopplung: Die kognitiven Anforderungen, die mit einer bestimmten Aufgabe verbunden sind, z. B. die Erkennung eines Signalabgriffs, hängt von der Verarbeitung der Informationen aus der Umgebung und der Ableitung der Beziehungen ab)
	PSF13	Fahrgefühl	fehlerhafte Einschätzung der Fahreigenschaften des Zuges
	PSF14	Fachwissen/Regelkenntnisse	Die Kenntnis und Anwendung von betrieblichen Regeln im Ereignisverlauf wirken dem Ereignisverlauf entgegen.

	PSF15	Streckenkenntnisse	eingehende Kenntnis der von einem Zug durchfahrenen Strecke.
	PSF16	Erfahrung	die Anhäufung von Wissen und Fähigkeiten, die durch die Interaktion mit dem System und die im Arbeitsumfeld verbrachte Zeit erworben werden
Job/System	PSF17	Aufgabenanforderung Zeitdruck	die Aufgabenanforderung in Bezug auf Anzahl und Art der Aufgaben sowie die für die Aufgaben zugewiesene Zeit
	PSF18	Kommunikation	der Informationsaustausch zwischen den Teammitgliedern über formale Kanäle und das gemeinsame Verständnis (Teamharmonie) der Situation
	PSF19	Kooperation	die Zusammenarbeit von am Ereignis beteiligten Personen
	PSF20	Unregelmäßigkeiten	Unbekanntheit mit einer Situation, die möglicherweise wichtig ist, aber nur selten auftritt oder die neu ist
	PSF21	Arbeitsplatzgestaltung	die Qualität des HMIs, ergonomisches Design, Anthropometrische Gestaltung
	PSF22	Lärm	interne/externe Arbeitsbedingungen - Arbeitsumweltbedingungen
	PSF23	Beleuchtung	interne/externe Arbeitsbedingungen - Arbeitsumweltbedingungen
	PSF24	Vibrationen	interne / externe Arbeitsbedingungen - Arbeitsumweltbedingungen
	PSF25	Wetterverhältnisse/Sichtverhältnisse	Einschränkung der Wahrnehmung sicherheitsrelevanter Informationen in der Außenansicht aufgrund von Wetterbedingungen
	PSF26	Ablöser	Einsatz als Ablöser
Organisational	PSF27	Qualität der Informationen	Mittel zur Übermittlung visueller oder akustischer Informationen/Rückmeldungen an den Arbeitnehmer zum richtigen Zeitpunkt in einer Form, die er leicht aufnehmen kann
	PSF28	Training/Weiterbildung	formales Training, um die erforderlichen Kenntnisse zu erwerben, um eine Arbeit genau und sicher ausführen zu können
	PSF29	Dienstplan	Personalplanung, Schichtmuster und Dauer, Schichtwechsel
	PSF30	Qualität der Verfahren, Anweisungen, Vorschriften	Vorhandensein und Vollständigkeit von Betriebsverfahren und anderen Referenzdokumenten für die Aufgabe oder die Situation

	PSF31	Sicherheitskultur	Werte, Einstellungen, Wahrnehmungen, Kompetenzen und Verhaltensmuster, die das Engagement sowie den Stil und die Kompetenz des Gesundheits- und Sicherheitsmanagements einer Organisation bestimmen (ACSNI, 1993); positive Sicherheitskultur, in der Informationen ohne Angst vor Diskriminierung kommuniziert werden
	PSF32	Aufsicht/Führungskultur	Vorhandensein und Einbeziehung von Koordinierungs-, Überwachungs-, Befehls- und Kontrollmaßnahmen
betriebliche/ technische Bedingungen	PSF33	Geschwindigkeit	Fahrgeschwindigkeit
	PSF34	Betriebsprogramm	Schienenpersonenfernverkehr (SPFV), Schienenpersonennahverkehr (SPNV), Schienengüterverkehr (SGV)
	PSF35	Gleisabstand	Der Gleisabstand ist der Abstand zwischen zwei nebeneinanderliegenden Gleisen, gemessen von Gleismitte zu Gleismitte.
	PSF36	Gleiszustand	Zustand der Gleise
	PSF37	Signalgestaltung	die Positionierung des Signals
	PSF38	Streckenauslegung	Profil der zu befahrenden Strecke
	PSF39	Blockabschnittslängen	die Länge des Blockabschnitts
	PSF40	Zugsicherungssystem	z. B. PZB, LCB/ETCS

LINK ANHANG 2: [PSF-EXPERTENSCHÄTZUNG \(.XLSX DATEI\)](#)

LINK ANHANG 3: [R-TOOL \(ZIP\)](#)