

Sachbericht zum Verwendungsnachweis

Teil II – Ausführlicher Bericht

WIR!-Projekt: Smart Sensor for Human Factors in Railway Systems
(HMI4Rail)
(FKZ 03WIR1214B)

**Zuwendungs-
empfänger:** Deutsches Zentrum für Luft- und Raumfahrt, Institut für
Verkehrssystemtechnik

Laufzeit: 03/2023 – 12/2025

Autorin: Dr. Esther Bosch

1 Durchgeführte Arbeiten und Ergebnisse

Im Folgenden werden die vom DLR Institut für Verkehrssystemtechnik durchgeführten Arbeiten und erzielten Ergebnisse in den relevanten Arbeitspaketen dargestellt. Die Arbeiten erfolgten in enger Abstimmung mit den Verbundpartnern der TU Chemnitz (SGT und ZWT) sowie dem Smart Rail Connectivity Campus (SRCC).

a) AP 1: Vitaldaten

AP 1.1 - Stand von Wissenschaft und Technik zur Detektion von Handlungsfähigkeit

In AP 1.1 wurde der grundlegende Forschungsstand zur sensorbasierten Detektion von Handlungsfähigkeit und Aufmerksamkeit von Bedienpersonal erarbeitet. Die Analyse zeigte, dass Fatigue zwar intensiv beforscht ist, eine einheitliche Definition jedoch bislang fehlt - bedingt durch die Vielschichtigkeit des Konzepts und die Wechselwirkung kognitiver, körperlicher und zirkadianer Einflussfaktoren. Als Arbeitsdefinition wurde mentale Ermüdung (Mental Fatigue, MF) als nicht-pathologischer, aufgabeninduzierter Zustand auf einem Kontinuum zwischen entspannter Wachheit und Schlaf verstanden, der kognitive Funktionen wie Daueraufmerksamkeit, Reaktionszeit und Handlungsüberwachung beeinträchtigt. Für die Vigilanzmessung gilt das EEG als Goldstandard. Diese Ausgangslage stellte eine wesentliche Orientierungsgröße für die Konzeption des Messsystems dar.

Im Bahnkontext ist MF insbesondere im Zusammenhang mit zunehmender Automatisierung und der damit einhergehenden Verlagerung der Triebfahrzeugführeraufgabe von aktiver Fahrzeugsteuerung hin zu passiver Überwachung (GoA2) sicherheitsrelevant. Verschiedene Parameter können aus Human-Factors-Sicht die Handlungsfähigkeit eines Triebfahrzeugführers beeinflussen: Vigilanz und Daueraufmerksamkeit bei Überwachungsaufgaben, Arbeitsbeanspruchung (Workload) sowie Ermüdung. Empirische Studien zeigen zudem, dass klassische Vigilanzüberwachungsgeräte (z. B. Sicherheitsfahrshalter (SIFA)) in ihrer Detektionsleistung begrenzt sind, da Triebfahrzeugführer hochautomatisierte Reaktionsmuster entwickeln können, die korrekte Systembedienung auch bei reduzierter Vigilanz ermöglichen.

Die Recherche identifizierte methodische Schwächen häufig eingesetzter Ansätze in der Literatur: Binäre Klassifikationsmodelle erhöhen zwar die rechnerische Klassifikationsgenauigkeit, sind für die praktische Echtzeit-Überwachung jedoch ungenügend, da gerade die graduellen Übergangphasen sicherheitsrelevant sind. Viele publizierte Studien basieren auf Kurzzeitlaborexperimenten mit geringer Stichprobengröße und unzureichender Kontrolle von Störgrößen. Daraus wurden für das Vorhaben folgende Anforderungen abgeleitet: aufgabennahes Design zur realitätsnahen Ermüdungsinduktion, Sicherstellung der Datenqualität durch Signalvalidierung und Ausreißerbehandlung, Einsatz multidimensionaler Referenzmaße sowie systematische Erfassung konfundierender Variablen.

MF wird typischerweise multimodal gemessen: mittels subjektiver Selbstberichte, (neuro-)physiologischer Signale sowie verhaltens- und leistungsbezogener Indikatoren. Während in der Automobilforschung und Luftfahrt bereits umfangreiche Erkenntnisse vorliegen und Multisensoransätze als vielversprechend gelten, ist die Evidenzbasis für den Bahnbereich deutlich schmaler. Insbesondere fehlen ökologisch valide Feldstudien unter realen Betriebsbedingungen. Diese Ausgangslage bildete die konzeptionelle Grundlage für die Versuchsplanung in AP 4.

AP 1.2 - Ableitung relevanter Vitalparameter und Anforderungen an das Datenerfassungssystem

Aufbauend auf AP 1.1 wurden in AP 1.2, in enger Abstimmung mit dem Projektpartner TU Chemnitz (SGT), konkrete Vitalparameter identifiziert und beschrieben, die zur Bewertung der Handlungsfähigkeit und Aufmerksamkeit von Triebfahrzeugführern geeignet sind. Die identifizierten Sensorsysteme wurden nach anatomischer Positionierung klassifiziert und umfassen kopf-, brust-, handgelenk-, ohr- und sprunggelenkgetragene Wearables,

Brillensysteme sowie Ganzkörperanzüge. Als meistuntersuchte Signalmodalitäten zur Ermüdungsdetektion wurden identifiziert: IMU, EEG, PPG, EKG, GSR und EMG; Hauttemperatur, Atemrate, EOG, NIRS sowie erweiterte Augenparameter gelten als ergänzend relevant, sind jedoch weniger etabliert. Ein zentraler praktischer Befund war, dass viele kommerzielle Systeme keinen offenen Rohdatenzugang bieten, was im Forschungskontext regelmäßig den Einsatz eigens evaluierter oder entwickelter Systeme erfordert.

Für den Einsatz im vorliegenden Projekt wurden folgende Messmodalitäten ausgewählt:

- Herzratenvariabilität (HRV): zeitdomänenbasierte Kennwerte (insbesondere RMSSD) als Indikator für autonome Aktivierung und parasympathische Aktivität
- Elektroenzephalographie (EEG): frontale Theta-Power als neurophysiologischer Marker kognitiver Ermüdung, ergänzt durch parietal Alpha- und Beta-Power sowie die Indizes Task Load Index (TLI) und Engagement Index (EI)
- Elektrodermale Aktivität (EDA/GSR): Hautleitwert-Peaks als Maß sympathischer Aktivierung
- Augenbewegungsparameter: Blinzel-Dauer als verhaltensbasierter Ermüdungsindikator
- Atemfrequenz: Verlangsamung als autonomes Korrelat reduzierter Arousal
- Subjektive Skalen: Karolinska Sleepiness Scale (KSS) als etabliertes Selbstberichtsmaß

Die Sensorauswahl orientierte sich sowohl an der theoretischen Fundierung als auch an den praktischen Anforderungen des Bahnbetriebs hinsichtlich Tragekomfort, Robustheit und Datenzugang. Diese Anforderungen bildeten die Grundlage für die Entwicklungsarbeiten der TU Chemnitz (SGT) in AP 3 sowie für die Versuchsplanung in AP 4.

b) AP 4: Versuchsplanung und -spezifikation (AP 4.2)

In AP 4.2 übernahm das DLR die Versuchsspezifikation und -planung für beide Studiumgebungen (Simulator und Realfahrt) in enger Abstimmung mit dem Projektpartner TU Chemnitz (SGT). Zentrales Ergebnis ist ein standardisiertes Versuchsdesign, das in beiden Umgebungen identisch umgesetzt wurde, um eine direkte Vergleichbarkeit der Ergebnisse zu gewährleisten.

Das Versuchsdesign umfasst folgende Kernelemente:

- Baseline-Fahrt (15 Minuten): Fahrt ohne zusätzliche kognitive Aufgabe als Referenzmessung
- Ermüdungsinduktionsparadigma (55 Minuten): Auditiver n-back-Task (PsychoPy, Version 2025.1.1) mit vier Schwierigkeitsstufen (0-back bis 3-back), jeweils zweimal fünf Minuten in randomisierter Reihenfolge; Eingabe der Zielzahl über Ziffernblock
- Post-Fatigue-Fahrt (15 Minuten): zweite Fahrt zur Erfassung der Ermüdungseffekte auf physiologische und Verhaltensmaße
- Übungsphase: Vor dem Experiment absolvierten alle Teilnehmenden eine Übungsphase des n-back-Tasks, bis eine Mindestgenauigkeit von 80 % auf allen Schwierigkeitsstufen erreicht wurde

Im Simulator erfolgte die Fahrt ohne PZB-System bei automatisch geregelter Geschwindigkeit; die Aufgabe der Teilnehmenden bestand im kontinuierlichen Bedienen der Sicherheitsfahrtschaltung (SIFA) sowie im Reagieren auf Streckenreize (Signalhorn bei Sicht auf orangefarbene Quader). In der Realfahrtstudie fuhren die Teilnehmenden unter GoA1-Bedingungen mit aktiver Geschwindigkeitsregelung und Bahnübergangsbeachtung. Das Design war in beiden Umgebungen so angelegt, dass es die Überwachungsaufgaben unter GoA2-Bedingungen so nah wie möglich abbildet.

Die Reaktionszeitaufgabe wurde durch das Platzieren von orangefarbenen 2x2-m-Würfeln (Simulator) bzw. Schneezäunen (Realfahrt) in 2 m Abstand zum Gleis operationalisiert; Teilnehmende sollten bei Sicht auf das Objekt das Signalhorn betätigen. Der Hintergrundkontrast wurde in beiden Umgebungen vergleichbar gestaltet.

Als zu messende Parameter und Bewertungskriterien wurden festgelegt: HRV (RMSSD, z-transformiert), Atemfrequenz (z-transformiert), EDA (GSR-Peaks), Blinzel-Dauer, EEG (frontale Theta-Power, parietal Alpha- und Beta-Power, Task Load Index, Engagement Index), KSS-Ratings, Reaktionszeiten auf Streckenobjekte sowie Anzahl verpasster SIFA-Auslösungen (nur Simulator).

Parallel zur Versuchsplanung entwickelte die TU Chemnitz (SGT) eine zentrale Messdatenbank, in die Datenströme aller eingesetzten Sensorsysteme einheitlich eingespeist wurden. Für jedes Messsystem wurde ein spezifischer Datenimporter erstellt, der die heterogenen Datenquellen in ein gemeinsames Format überführt. Die im Rahmen der Simulator- und Realfahrten erhobenen DLR-Datensätze wurden nach der Versuchsdurchführung an die TU Chemnitz übermittelt, dort aufbereitet und in die Datenbank eingepflegt; der DLR-seitige Zugriff auf vorverarbeitete Daten erfolgte über einen eingerichteten Fernzugriff. Eine direkte Anbindung bereits während der Messfahrten hätte ausgefallene Sensoren unmittelbar sichtbar machen können; dies ist ein Aspekt, der bei künftigen Versuchskampagnen berücksichtigt werden sollte.

c) **AP 6: Umsetzung Simulatorversuche**

AP 6.2 – Durchführung der Simulatorversuche

Die Simulatorstudie wurde am DLR Institut für Verkehrssystemtechnik in Braunschweig im Zeitraum Juli bis August 2024 durchgeführt. Der Ethikantrag wurde durch das DLR eingereicht und genehmigt. Die Versuchsinfrastruktur basiert auf der Simulationsumgebung RailSET® (Railway Simulation Environment for Train Drivers and Operators), die über ein originales Steuerungspult eines Triebfahrzeuges, eine projizierte Vorwärtsansicht und Seitenbildschirme sowie ein vollständiges Kabinenaudio-System verfügt.

An der Studie nahmen 14 professionelle Triebfahrzeugführer teil (13 männlich, 1 weiblich; mittleres Alter: $36,5 \pm 15,4$ Jahre; mittlere Berufserfahrung: 14 ± 15 Jahre). Alle Teilnehmenden gaben schriftliches Einverständnis und erhielten eine Aufwandsentschädigung von 35 EUR pro Stunde. Die Gesamtdauer pro Versuchsperson betrug ca. zwei Stunden.

Die physiologische Sensorausstattung umfasste:

- EEG: Brain Products LiveAmp-32 mit aktiven Ag/AgCl-Elektroden (gelbasiert), 26 Elektroden im 10-20-System (BrainVision Recorder 1.26, 500 Hz)
- Eye-Tracking: Pupil Core Brille zur Erfassung von Blinzel-Dauer
- EKG/HRV: Polar H10 Brustgurt
- Atemfrequenz: Vernier Go Direct Atemgurt
- Hautleitfähigkeit: Shimmer3 GSR-Sensor
- Subjektives Maß: Karolinska Sleepiness Scale (KSS) vor der ersten sowie nach jeder Fahrt

AP 6.5 / 6.6 – Vergleich und Diskussion der Ergebnisse

Die Auswertung der Simulatordaten erfolgte sensorspezifisch (siehe Abbildung 1). Da die Stichprobengrößen keine inferenzstatistischen Analysen mit ausreichender Teststärke erlauben, wurden die Ergebnisse primär deskriptiv berichtet und theoretisch eingeordnet. Die zentralen Befunde sind:

- HRV (RMSSD): konsistentes Muster in erwarteter Richtung; niedrigere Werte in Fahrt 1, höhere Werte in Fahrt 3 (Simulator: $M = -0,35$ vs. $0,32$, z-transformiert), was auf erhöhte parasympathische Aktivität und reduzierten Arousal nach der Ermüdungsinduktion hinweist
- Atemfrequenz: komplementäres Muster; leicht erhöhte Werte in Fahrt 1, reduzierte Werte in Fahrt 3, konsistent über beide Umgebungen
- Subjektive Karolinska Sleepiness Scale (KSS): deutlicher Anstieg in beiden Umgebungen ($M = 3,21$ vor der n-back-Aufgabe bis $M = 5,50$ nach der Aufgabe), was auf eine erfolgreiche Ermüdungsinduktion deutet

- EEG: frontale Theta-Power zeigte keine hypothesenkonforme Zunahme; Theta-Aktivität nahm an den Elektroden F3 und F4 zwischen den Fahrten ab, was den Befunden von Hamann und Carstengerdes (2023) aus einer Pilotenstudie widerspricht
- EDA (GSR-Peaks): kein bedeutsamer Unterschied zwischen den Fahrten im Simulator (M = 40,6 vs. 41,4 Peaks), hohe interindividuelle Variabilität
- Blinzel-Dauer: durchgängig negative z-Werte (unterhalb der individuellen Baseline), nahezu identisch zwischen Fahrt 1 und Fahrt 3, kein differenzieller Ermüdungseffekt
- Reaktionszeiten auf Streckenobjekte: kürzere Reaktionszeiten in Fahrt 3 (M = 3,68 s) als in Fahrt 1 (M = 4,51 s), was auf einen Übungseffekt hindeutet
- Verpasste SIFA-Auslösungen: sehr niedrige Frequenz und kein bedeutsamer Unterschied zwischen den Fahrten, was auf die bekannte Limitation klassischer Vigilanzüberwachungsgeräte hinweist

Die Ergebnisse wurden auf der 6th German Conference on Rail Human Factors (Berlin, Februar 2025) vorgestellt und als Konferenzbeitrag veröffentlicht (Bosch et al., 2025, DOI: 10.82350/X4EC-CH40). Ein erweiterter Beitrag mit Einschluss der Realfahrt Daten wurde auf der 8th International Conference on Rail Human Factors (London, September 2025) veröffentlicht.

d) AP 5: Umsetzung realer Testszenerien

Die Realfahrtstudie wurde am 8. und 9. März 2025 in Zusammenarbeit mit dem Smart Rail Connectivity Campus (SRCC) und der TU Chemnitz auf einer gesperrten Teststrecke im Bereich Schwarzenberg durchgeführt. Das DLR verantwortete das wissenschaftliche Studiendesign sowie die Erhebung und Auswertung der physiologischen und verhaltensbezogenen Daten; die organisatorische und infrastrukturelle Umsetzung vor Ort erfolgte in enger Abstimmung mit dem SRCC. Die Strecke war für die gesamte Versuchsdauer für den regulären Betrieb gesperrt.

An der Realfahrtstudie nahmen sechs professionelle Triebfahrzeugführer teil (alle männlich; mittlere Berufserfahrung: 23 ± 2 Jahre). Das Versuchsdesign entsprach dem der Simulatorstudie. Die physiologische Sensorausstattung war weitgehend identisch; für das EEG wurde ein AntNeuro eego sports (64 Schwammelektroden, 500 Hz) eingesetzt.

Ergebnisse der Realfahrtstudie

Die Realfahrt Daten replizieren die zentralen Befunde der Simulatorstudie für die autonomen Indikatoren und ergänzen sie um feldspezifische Erkenntnisse:

- HRV (RMSSD): konsistentes Muster analog zur Simulatorstudie; Fahrt 1 M = -0,31 (SD = 0,40), Fahrt 3 M = 0,32 (SD = 0,42, z-transformiert), was auf erhöhte parasympathische Aktivität nach der Ermüdungsinduktion hindeutet. Die Übereinstimmung mit dem Simulatorbefund stärkt die Verlässlichkeit dieses Indikators über Umgebungen hinweg.
- Atemfrequenz: paralleles Muster; Fahrt 1 M = 0,13 (SD = 0,15), Fahrt 3 M = -0,11 (SD = 0,16, z-transformiert), was die Befunde aus dem Simulator repliziert und die ökologische Robustheit der Atemrate als Ermüdungsindikator unterstreicht.
- KSS: geringerer Anstieg der subjektiven Müdigkeit als im Simulator (Fahrt 1 M = 2,17 vs. Fahrt 3 M = 3,33), was möglicherweise auf höhere motivationale Aktivierung unter realen Betriebsbedingungen zurückzuführen ist.
- EDA (GSR-Peaks): leichter Anstieg von Fahrt 1 (M = 46,2, SD = 17,85) zu Fahrt 3 (M = 58,6, SD = 22,33), was der Hypothese reduzierter sympathischer Aktivierung widerspricht. Bei n = 5 und hoher interindividueller Variabilität ist dieser Befund jedoch mit Vorsicht zu interpretieren.
- Blinzel-Dauer: nahezu identische z-Werte zwischen Fahrt 1 (M = -0,70, SD = 0,14) und Fahrt 3 (M = -0,70, SD = 0,13), kein ermüdungsrelevanter Effekt. Blinzel-Dauern waren im Realfahrtkontext negativer als im Simulator, was auf höhere visuelle Aufmerksamkeit unter realen Fahrbedingungen hindeutet.

- Reaktionszeiten: deutlicher Rückgang von Fahrt 1 (M = 11,65 s, SD = 6,01) zu Fahrt 3 (M = 3,46 s, SD = 3,90) bei n = 3, was vorrangig auf einen Übungseffekt zurückzuführen ist.
- EEG: Die Auswertung der Realfahrt-EEG-Daten war durch substantielle technische Herausforderungen beeinträchtigt (s. u.); verwertbare Daten lagen nur für einen Teil der Teilnehmenden vor.

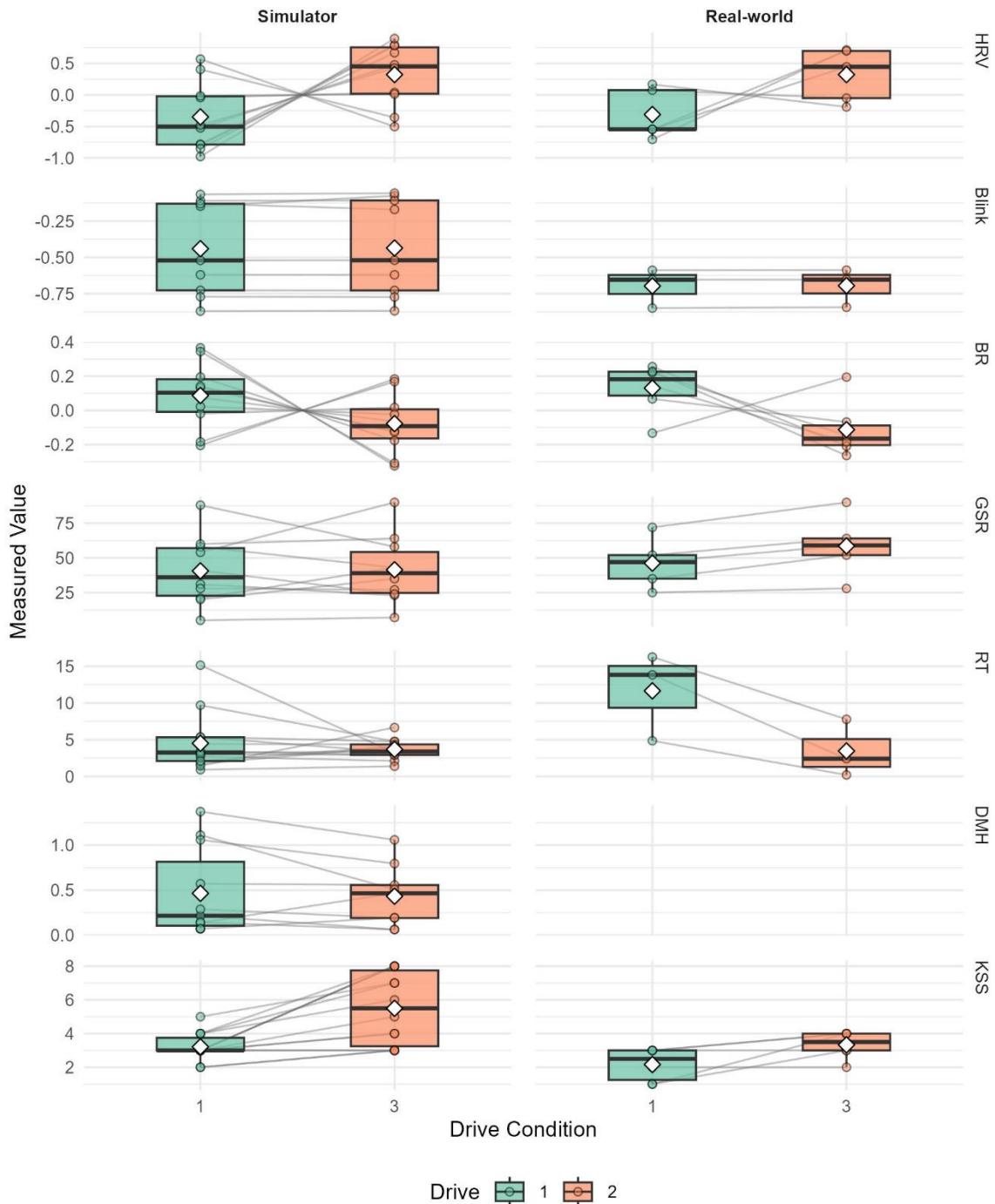


Abbildung 1: deskriptive Statistik der verschiedenen MF-Maße in Simulator und Realwelt-Umgebung.

Technische Herausforderungen im Feldeinsatz

Die Realfahrtdurchführung offenbarte mehrere feldspezifische Herausforderungen, die für die Weiterentwicklung operativer Messsysteme im Bahnbereich von zentraler Bedeutung sind:

- USB-Verbindungen erwiesen sich durch Fahrzeugvibrationen als kritische Schwachstelle; Wackelkontakte führten zu Datenverlust bei Eye-Tracking und Kamerasystem
- Gleichzeitiger Betrieb mehrerer Hochfrequenz-Bluetooth-Sensoren (GSR, EKG, Atemgurt, EEG, Reaktionstaste) führte zu Überlastung der Drahtlosübertragung
- Die mobile Energieversorgung (keine stationären Netzanschlüsse im Zug) erforderte ein eigens entwickeltes Konzept mit ausreichender Kapazität für alle Messkomponenten

Diese Erkenntnisse unterstreichen, dass ein direkter Transfer laborbasierter Sensoraufbauten in den Feldeinsatz ohne systematisches Robustheits-testing nicht zielführend ist, und fließen in die Anforderungsableitung in AP 7 ein.

e) AP 7: Ableitung von Anforderungen an technische Unterstützungssysteme (AP 7.7)

Auf Basis der zusammengeführten Ergebnisse aus Simulator- und Realfahrtstudie wurden konkrete Anforderungen an technische Unterstützungssysteme für Triebfahrzeugführer abgeleitet.

Als Gesamtbefund lässt sich festhalten: Kein einzelner Indikator bietet eine umfassende oder universell zuverlässige Messung mentaler Ermüdung unter Betriebsbedingungen. HRV und Atemfrequenz erwiesen sich als die ökologisch robustesten Indikatoren, da sie in beiden Umgebungen konsistente und theoretisch erwartete Veränderungen zeigten, relativ wenig artefaktanfällig sind und sich für die Integration in Brustsensoren und Smart Garments eignen. EEG zeigte trotz starker Laborevidenz keine konsistenten ermüdungsrelevanten Muster und ist unter Feldbedingungen mit erheblichen praktischen Herausforderungen verbunden.

Abgeleitete Anforderungen an ein praxistaugliches Echtzeit-Überwachungssystem:

- Minimalinvasivität: Das System soll den Triebfahrzeugführer so wenig wie möglich in seiner Arbeit beeinträchtigen; Bevorzugung körpernaher, unauffälliger Sensorik (z. B. Brustgurt, Smart Garment)
- Robustheit gegenüber Vibrationen und Bewegungsartefakten: Kabelgebundene USB-Verbindungen sind durch vibrationsstabile Alternativen zu ersetzen; drahtlose Hochfrequenzsensorik erfordert dedizierte Übertragungsarchitektur
- Priorisierung autonomer Indikatoren: HRV (RMSSD) und Atemfrequenz sind als Kernindikatoren für ein operatives System zu priorisieren
- Ökologisch valide Ermüdungsinduktion: Zukünftige Validierungsstudien sollten passive Ermüdung durch monotone Überwachungsaufgaben induzieren, die der realen GoA2-Situation näherkommen, anstelle aktiver kognitiver Ermüdung durch n-back-Aufgaben
- Datenschutz und Akzeptanz: Das System soll das Monitoring auf die Dienstzeit beschränken; Datensicherheit ist als Zentralkriterium zu berücksichtigen; Stakeholder-Rückmeldungen (Digital Rail Convention 2023) unterstreichen die Bedeutung dieser Anforderungen
- Stichprobengröße: Replikationsstudien erfordern größere Stichproben für inferenzstatistisch belastbare Aussagen zu Umgebungsunterschieden und Sensorleistung

Auf Basis der zusammengeführten Ergebnisse aus Simulator- und Realfahrtstudie wurden initiale Orientierungswerte für ausgewählte Belastungsindikatoren abgeleitet. Diese sind als vorläufig zu betrachten und bedürfen der Validierung anhand größerer Datensätze unter variierenden Realbedingungen. Als übergreifendes empirisches Ergebnis zeigte sich, dass Herzratenvariabilität und Atemfrequenz die konsistentesten und ökologisch robustesten Muster lieferten; beide Indikatoren werden daher als pragmatischer Ausgangspunkt für ein operatives Zustandsüberwachungssystem empfohlen.

Ergänzende Erkenntnisse für die Anforderungsableitung lieferte die Stakeholder-Veranstaltung im Rahmen der Digital Rail Convention 2023 in Annaberg-Buchholz, an der Vertreterinnen und Vertreter des DZSF, der DB Netz AG, der IHK Chemnitz sowie der IAV GmbH teilnahmen. Die

Diskussion ergab, dass ein praxistaugliches System zur Zustandsüberwachung minimalinvasiv und auf das Wesentliche reduziert sein sollte, Datensicherheit als zentrales Kriterium berücksichtigen und das Monitoring auf die Arbeitszeit beschränken sollte. Diese Rückmeldungen schärften die Anforderungen an den Demonstrator und fließen in die Empfehlungen für Folgestudien ein.

2 Wichtigste Positionen des zahlenmäßigen Nachweises

Die wesentlichen Ausgaben entfielen auf Personalkosten für die wissenschaftliche Bearbeitung der Arbeitspakete AP 1, AP 4, AP 6 und AP 7 durch Mitarbeiterinnen und Mitarbeiter des DLR Instituts für Verkehrssystemtechnik. Reisekosten entstanden für die Beteiligung an Verbundtreffen sowie die Präsentation der Projektergebnisse auf Fachkonferenzen. Sachkosten umfassten Kosten für die Erstellung der Streckenvisualisierung sowie den Einsatz der Großanlage RailSET®. Wesentliche Abweichungen vom Finanzplan liegen nicht vor.

3 Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Die durchgeführten Arbeiten entsprechen dem im Zuwendungsbescheid beschriebenen Arbeitsplan und waren zur Erreichung der definierten Projektziele notwendig. Das DLR verfügt über die einzigartige Infrastruktur des RailSET®-Simulators sowie über langjährige Expertise in der Human-Factors-Forschung im Bahnbereich, die für die Durchführung der Simulatorstudie unabdingbar war. Die Arbeiten in AP 1 und AP 4 bildeten die konzeptionelle Grundlage für alle empirischen Arbeiten des Verbundes. Die Auswertung und Ergebnisdiskussion in AP 6 und AP 7 sichern die wissenschaftliche Verwertbarkeit der erhobenen Daten. Der Personalmitteleinsatz war für die Komplexität und den Umfang der durchgeführten Arbeiten angemessen.

4 Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse

Die Projektergebnisse sind auf mehreren Ebenen verwertbar:

Wissenschaftlich liefert das Vorhaben erstmals empirisch fundierte, ökologisch valide Evidenz zur Eignung physiologischer Indikatoren für die Ermüdungsüberwachung von Triebfahrzeugführern sowohl im Simulator als auch unter Realbedingungen. Die Nichtreplikation erwarteter EEG-Befunde aus der Luftfahrtforschung ist ein eigenständig bedeutsamer Befund, der die Kontextsensitivität neurophysiologischer Ermüdungsmarker unterstreicht und die Notwendigkeit domänenspezifischer Validierungsstudien hervorhebt. Die Ergebnisse wurden auf der 6th German Conference on Rail Human Factors (2025) und der 8th International Conference on Rail Human Factors (London, September 2025) publiziert. Eine Einreichung bei einer internationalen Fachzeitschrift (Applied Ergonomics) ist in Vorbereitung.

Methodisch stellen das entwickelte Versuchsdesign und die standardisierte Sensorausstattung wiederverwendbare Infrastruktur für Folgestudien zur Verfügung. Die im Projekt gewonnenen Erkenntnisse zu feldtauglicher Datenerhebung (Vibrationsproblematik, Bluetooth-Saturation, Energieversorgung) sind für zukünftige Feldstudien im Bahnbereich von unmittelbarem praktischem Wert.

Für die Systementwicklung bilden die abgeleiteten Anforderungen (AP 7.7) eine konzeptionelle Grundlage für die Entwicklung von Fahrerzustandsüberwachungssystemen, Smart Garments und mensch-zentrierten Automatisierungsstrategien im Bahnbetrieb. In Verbindung mit den Kompetenzen der beteiligten regionalen KMU besteht mittelfristig Potenzial für industrielle Anschlussprojekte.

Das Vorhaben stärkt die Position des DLR Instituts für Verkehrssystemtechnik als eines der wenigen deutschen Forschungsinstitute, das Human-Factors-Forschung im Bahnbereich mit physiologischer Mehrkanalmesstechnik verbindet. Die gewonnenen Erkenntnisse und der entwickelte Versuchsaufbau bilden die Grundlage für ein geplantes Folgeprojekt (HMI4Rail²), das die Thematik der Zustandsüberwachung von Bahnpersonal in einen erweiterten Kontext kommunikationsbasierter Betriebssicherung überführt.

5 Während der Durchführung bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Im Projektzeitraum ist der wissenschaftliche Fortschritt auf dem Gebiet der Zustandsüberwachung von Triebfahrzeugführern weiter vorangeschritten. Der erkennbare Trend in der internationalen Forschung geht zunehmend in Richtung kamerabasierter, KI-gestützter Systeme zur Ermüdungsdetektion. Ansätze auf Basis von Graph Neural Networks kombinieren Gesichts- und Skelettmerkmale zur Klassifikation von Aufmerksamkeitszuständen bei Triebfahrzeugführern (vgl. Nocentini et al., 2025). Dieser methodische Zugang ist komplementär zum wearable-basierten Monitoring: Während kamerabasierte Systeme keinen Körperkontakt erfordern, bieten physiologische Wearables höhere Signalspezifität und Unabhängigkeit von Lichtverhältnissen und Verdeckungen im Führerstand. Eine Kombination beider Ansätze erscheint mittelfristig vielversprechend.

Im deutschen Förderkontext sind dem Projektteam keine direkt vergleichbaren laufenden Vorhaben bekannt, die denselben Ansatz eines körpergetragenen Multisensor-Systems zur Vigilanzüberwachung im Bahnbetrieb verfolgen. HMI4Rail besetzt damit eine in Deutschland bislang weitgehend unbesetzte Forschungsnische, insbesondere durch die Kombination aus Simulatorstudie, Realfahrtvalidierung und dem systematischen Vergleich physiologischer Indikatoren unter operativen Bedingungen.

6 Erfolgte oder geplante Veröffentlichungen

Folgende Veröffentlichungen sind aus dem Projekt hervorgegangen oder befinden sich in Vorbereitung:

- Bosch, E. J., Schackmann, D., Hoyer, S., Kilian, W., Schwanitz, S., & Hamann, A. (2025). Multi-Sensor Vigilance Detection: Sensor Reliability and Comfort in Automated Train Operations. *6th German Conference on Rail Human Factors*, Berlin, 18. Februar 2025. <https://doi.org/10.82350/X4EC-CH40>
- Bosch, E. J., Hoyer, S., Kilian, W., Schwanitz, S., Schackmann, D., & Hamann, A. (2025). Multi-Sensor Vigilance Detection in Train Drivers – a Simulator Study. *8th International Conference on Rail Human Factors*, London, 18. September 2025.
- Bosch, E. J. et al. (in Vorbereitung). Multisensor Measurement of Train Driver Fatigue: Lessons Learned. *Applied Ergonomics* (geplante Einreichung 2026).

Referenzen

Hamann, A., & Carstengerdes, N. (2023). Assessing the development of mental fatigue during simulated flights with concurrent EEG-fNIRS measurement. *Scientific Reports*, 13(1), 4738.

Nocentini, O., Lagomarsino, M., Solak, G., Cho, Y., Tong, Q., Lorenzini, M., & Ajoudani, A. (2025). Graph-based online monitoring of train driver states via facial and skeletal features. *arXiv preprint arXiv:2505.08800*.