

Fachbereich Wirtschaftswissenschaft

**Entwicklung und Evaluation intelligenter Systeme in Leitstellen der
Behörden und Organisationen mit Sicherheitsaufgaben aus
Perspektive der Endanwender:innen**

Dissertation

zur Erlangung der Doktorwürde

durch den

Promotionsausschuss Dr. rer. pol.

der Universität Bremen

vorgelegt von

Christian Elsenbast, M.A., MSc, MA

Hochspeyer, 03.11.2024

Gutachter:innen:

Erstgutachterin Frau Prof. Dr. Vera Hagemann

Zweitgutachterin Frau Prof. Dr. Julia Kensbock

Dissertationskolloquium:

Bremen, 18.03.2025

I Abstract Deutsch

Einleitung: Leitstellen der Behörden und Organisationen mit Sicherheitsaufgaben (BOS) stellen komplexe Arbeitsumgebungen dar. Die Leitstellendisponent:innen arbeiten oftmals bei unvollständigen Informationen unter Zeitdruck, um weitreichende und unumkehrbare Entscheidungen zu treffen. Basierend auf empirischen Erkenntnissen zielt diese Dissertation darauf ab, Empfehlungen und Hinweise zur Entwicklung und Evaluation von intelligenten Systemen, die den Leitstellendisponent:innen als Entscheidungsunterstützungssysteme (z. B. bei der Alarmierung der Rettungsmittel) dienen, aus Perspektive der Endanwender:innen zu geben.

Methoden: Zur Zielerreichung wurden zwei Fragebogenstudien zu soziodemografischen Parametern, Technikbereitschaft und Team-Arbeit-Kontexten entwickelt und online an die Leitstellendisponent:innen und an weitere relevante Berufsgruppen (z. B. Notfallsanitäter:innen) adressiert. Außerdem wurden im Rahmen der Grounded-Theory-Methodology Leitstellendisponent:innen, Führungskräfte und andere Expert:innen (z. B. Informatiker:innen) zu Überzeugungssystemen gegenüber intelligenten Systemen in Leitstellen interviewt. Um Forschungslücken zu den Arbeitsprozessen der Leitstellendisponent:innen zu schließen, wurde eine Eyetracking-Studie am Leitstellenarbeitsplatz durchgeführt.

Ergebnisse: Die Studie zur Technikbereitschaft (N = 510) deutet auf eine grundsätzliche gute Technikbereitschaft hin, auch gegenüber intelligenten Systemen. Jedoch waren auch Personen mit sehr geringer Technikbereitschaft als statistische Ausreißer zu verzeichnen. Die Studie zum Team-Arbeit-Kontext (N = 490) zeigt u. a., dass die Leitstellendisponent:innen ihre Tätigkeit als komplex wahrnehmen und ermöglicht die Identifikation einzelner komplexitätsbringender Faktoren (z. B. Intransparenz). Außerdem konnte durch 31 Personen, die über 619 Minuten interviewt wurden, per Grounded Theory herausgestellt werden, dass Qualifikationsniveau, beruflicher Status und Vorwissen zu künstlicher Intelligenz Prädiktoren für eine positive Haltung gegenüber intelligenten Systemen in Leitstellen sind. Weiterhin wurde gezeigt, dass Unterstützungsbedarf bei Visualisierung und Automatisierung bestehen. Die Eyetracking-Studie (N = 33) konnte Areale hoher und geringer visueller Aufmerksamkeit darstellen und weiterhin praxisrelevante Unterschiede zwischen einem Vier- und Zwei-

Monitor-Setting darstellen. Außerdem konnten 5 Evaluationsphasen und 11 Entwicklungsschritte für ein intelligentes System expliziert werden.

Diskussion: Hinsichtlich intelligenter Systeme sind die Leitstellendisponent:innen aufgeschlossen, jedoch dürfen Menschen mit geringer Technikbereitschaft nicht vergessen werden. Maßnahmen zur Komplexitätsreduktion sind als Entwicklungsziel bei intelligenten Systemen zu berücksichtigen. Bei der Entwicklung sind u. a. Feedback an die Nutzer:innen, kontextsensitive Visualisierungen und Automatisierung von Standardaufgaben zu berücksichtigen. Die intelligenten Systeme sind durch ein individuell abgestimmtes Bündel an organisatorischen, persönlichen und verhaltensbezogenen Maßnahmen zu ergänzen, die verbindliche Standards fokussieren. Zur leitliniengerechten Evaluation der intelligenten Systeme dienen die vorgestellten Evaluationsphasen. Weitere Studien, die größere Subgruppenvergleiche ermöglichen, sowie Eyetracking-Studien, die verschiedene technische Set-ups der Leitstellen berücksichtigen und somit bessere Vergleichbarkeit der Aufgaben gewährleisten, sind wünschenswert.

II Abstract English

Introduction: Public Safety Answering Points (PSAPs) of authorities and organisations with security responsibilities (AOS) represent complex working environments. PSAP dispatchers often work with incomplete information under time pressure to make far-reaching and irreversible decisions. Based on empirical findings, this dissertation aims to provide recommendations and guidance for the development and evaluation of intelligent systems in AOS PSAPs from the end-users' perspective. These intelligent systems serve as decision support systems for dispatchers, e.g., in alerting emergency resources.

Methods: To achieve these objectives, two questionnaire studies on socio-demographic parameters, technology readiness, and team work contexts were developed and distributed online to dispatchers and other relevant professional groups (e.g., paramedics). Furthermore, within the framework of Grounded Theory Methodology, dispatchers, managers, and other experts (e.g., computer scientists) were interviewed about belief systems regarding intelligent systems in PSAPs. To address research gaps concerning dispatchers' work processes, an eye-tracking study was conducted at the PSAP workplace.

Results: The technology commitment study (N = 510) indicates a generally good level of technology readiness, including towards intelligent systems. However, some individuals with very low technology commitment were recorded as statistical outliers. The team work context study (N = 490) shows, among other findings, that dispatchers perceive their work as complex and enables the identification of specific complexity-inducing factors (e.g., lack of transparency). Moreover, through interviews with 31 individuals conducted over 619 minutes, Grounded Theory analysis revealed that qualification level, professional status, and prior knowledge of artificial intelligence are predictors of a positive attitude towards intelligent systems in PSAPs. Furthermore, it was shown that support is needed in visualisation and automation. The eye-tracking study (N = 33) was able to identify areas of high and low visual attention and demonstrate differences between four- and two-monitor settings. Additionally, 5 evaluation phases and 11 development steps for an intelligent system were explicated.

Discussion: Regarding intelligent systems, dispatchers are generally open-minded; however, individuals with low technology readiness must not be overlooked. Measures for complexity reduction should be considered as a development goal for intelligent systems. During development, user feedback, context-sensitive visualisations, and automation of standard

tasks must be taken into account. The intelligent systems should be complemented by an individually tailored bundle of organisational, personal, and behavioural measures that place binding standards at the centre of consideration. The presented evaluation phases serve for guideline-compliant evaluation of intelligent systems. Further studies enabling larger subgroup comparisons, as well as eye-tracking studies considering various technical PSAP set-ups and thus ensuring better comparability of tasks, are desirable.

III Inhaltsverzeichnis

I Abstract Deutsch	ii
II Abstract Englisch	iv
III Inhaltsverzeichnis	vi
IV Tabellenverzeichnis	ii
V Abbildungsverzeichnis	iii
VI Abkürzungsverzeichnis	iv
VII Publikationen im Zusammenhang mit der Dissertation	v
1 Einleitung und Relevanz des Forschungsthemas	1
1.1 Heranführung	1
1.2 Technik- und menschenzentrierte Perspektiven auf KI in den Leitstellen	2
1.3 Ziele und Forschungsfragen	4
2 Komplexität in Leitstellen	6
3 Methodologischer Rahmen	9
3.1 Verortung des Forschungshintergrundes mit methodologischen Erwägungen	9
3.2 Zentrale Begriffe und Konzepte	10
3.2.1 Technikbereitschaft	10
3.2.2 Team-Arbeit-Kontext	12
3.2.3 Überzeugungssysteme	12
3.2.4 Visuelle Aufmerksamkeit	13
3.3 Zentrale Methoden und methodologische Erwägungen	14
3.3.1 Fragebogenforschung und Psychometrie	14
3.3.2 Interviews und Grounded-Theory-Methodology	15
3.3.3 Eyetracking	17
4 Begründung und Konkretisierung von Forschungsfragen, Forschungsdesign, Zuordnung von Forschungsfragen zu Publikationen und Zusammenfassung der Publikationen	19
4.1 Begründung der Forschungsfragen	19
4.1.1 RQ 1 – Demografie, Technikbereitschaft und Team-Arbeit-Kontext	19
4.1.2 RQ 2 – Überzeugungssystem der Leitstellendisponent:innen	20
4.1.3 RQ 3 – Aufmerksamkeitsverteilung der Disponent:innen beim Notrufdialog	20
4.2 Konkretisierung, Begründung und wissenschaftstheoretische Verortung des Forschungsdesigns	22
4.2.1 RQ 1 – Demografie, Technikbereitschaft und Team-Arbeit-Kontext	22
4.2.2 RQ 2 – Überzeugungssystem der Leitstellendisponent:innen	23
4.2.3 RQ 3 – Aufmerksamkeitsverteilung der Disponent:innen beim Notrufdialog	24
4.3 Zuordnung der Publikationen zu den Forschungsfragen	25
4.4 Zusammenfassung der Publikationen	26
4.4.1 P1 - Technology commitment of emergency medical service practitioners and dispatchers	26
4.4.2 P2 und P3 - Team-Arbeit-Kontext	27
4.4.3 P4 - Mental Models of AI in Emergency Medical Services	29

4.4.4 P5 - Towards Human-AI Interaction in Medical Emergency Call Handling	31
4.4.5 P6 – Visuelle Aufmerksamkeit in Leitstellen der Gefahrenabwehr	32
5 Diskussion.....	34
5.1 Interpretation	34
5.1.1 RQ 1 - Soziodemografie, Technikbereitschaft und Team-Arbeit-Kontext.....	34
5.1.2 RQ 2 - Überzeugungssystem der Leitstellendisponent:innen	46
5.1.3 RQ 3 - Aufmerksamkeitsverteilung der Disponent:innen beim Notrufdialog	49
5.2 Ausblick und weitere Forschung.....	54
5.3 Fazit.....	55
Literaturverzeichnis.....	57
Anlage 1 - Eigenständigkeitserklärung	a
Anlage 2 – Eigenleistungen	b
Anlage 3 – Elaboration Expert:innen-Interview	c
Anlage 4 – Ergebnistabelle TAKAI (P2 und P3).....	e
Anlage 5 – Transfer der STOPV-Maßnahmen	g
Anlage 6 – Explainable AI – Umsetzungsimpulse.....	h
Anlage 7 - Ethische Exkursion.....	i
Anlage 8 – Tabellarische Zusammenfassung der Publikationen	k
Anlage 9 – Publikationen.....	p

IV Tabellenverzeichnis

Tabelle 1: Liste der dissertationsrelevanten Publikationen mit Peer-Review-Verfahren	v
Tabelle 2: Gütekriterien qualitativer Forschung (nach Steinke, 2019, S. 319-331)	17
Tabelle 3: Zuordnung der Publikationen zu den Forschungsfragen	25
Tabelle 4: Beschreibung der Eigenleistungen	b
Tabelle 5: Ergebnisse der Itemgruppen und Intraklassen-Korrelation	e
Tabelle 6: STOPV-Maßnahmen	g
Tabelle 7: Explainable AI - Umsetzungsimpulse.....	h
Tabelle 8: Tabellarische Zusammenfassung der Publikationen	k

V Abbildungsverzeichnis

Abbildung 1: Stufen der Automatisierung angelehnt an Parasuraman (2020)	10
Abbildung 2: Eyetracking Set-up in der Leitstelle	19
Abbildung 3: Boxplots - Technikbereitschaft	27

VI Abkürzungsverzeichnis

BMWK	Bundeministerium für Wirtschaft und Klimaschutz
BOS	Behörden- und Organisationen mit Sicherheitsaufgaben
CDS	Calltaker-Dispatcher-Systeme
CFA	Confirmatory factor analysis (Konfirmatorische Faktorenanalyse)
CIRS	Critical Incident Reporting System
CLT	Cognitive Load Theory (Theorie der kognitiven Belastung)
DSS	Decision Support System (Entscheidungsunterstützungssystem)
ELS	Einsatzleitsystem
ESS	Einsatzsachbearbeiter-System
EMH	Eye-Mind Hypothese
GTM	Grounded-Theory-Methodology
HRT(s)	High-Responsibility-Team ¹
ICC	Intraclass correlation (Intraklassen-Korrelation)
ILS	Integrierte Leitstellen für Brandschutz, Katastrophenschutz und Rettungsdienst
KI	Künstliche Intelligenz
KRITIS	Kritische Infrastrukturen
NotSan	Notfallsanitäter:innen
P	Publikation
RQ	Research Question (Forschungsfrage)
SNA	Standardisierte Notrufabfrage
SPELL	Semantische Plattform zur intelligenten Entscheidungs- und Einsatzunterstützung in und beim Lagemanagement
TAKAI	Team-Arbeit-Kontext-Analyse Inventar
THW	Technisches Hilfswerk
TRM	Team-Ressource-Management
XAI	Explainable Artificial Intelligence (Erklärbare künstliche Intelligenz)
ZNA	Zentrale Notaufnahme

¹ Da eine deutsche Entsprechung nicht gebräuchlich ist, wird auf eine Übersetzung verzichtet.

VII Publikationen im Zusammenhang mit der Dissertation

Tabelle 1: Liste der dissertationsrelevanten Publikationen mit Peer-Review-Verfahren

Nr.	RQ	Titel	Journal (J) /Konferenz (K)	Autor:innen	Status
P1	1	Technology commitment of emergency medical service practitioners and dispatchers	International Paramedic Practice (J)	Elsenbast, C. & Hagemann, V.	Veröffentlicht
P2		Komplexität und Adaptionserfordernisse an der Nahtstelle Leitstelle-Rettungsdienst-Feuerwehr: Eine validierte Fragebogenanalyse	German Journal of Paramedic Science (J)	Elsenbast, C.	Im Review
P3		Zwischen Hierarchie und geteilten mentalen Modellen: Analyse mentaler Modelle an der Nahtstelle Leitstelle-Rettungsdienst-Feuerwehr			
P4	2	Attitude of Emergency Dispatchers Towards Artificial Intelligence - A Black Box of Expectations	International Journal Paramedicine (J)	Elsenbast, C., Reuter-Oppermann, M. & Philipp Dahlmann	Im Review
P5	3	Towards Human-AI Interaction in Medical Emergency Call Handling	HICCS, Hawaii International Conference on System Sciences, 2024 (K)	Maletzki, C., Elsenbast, C. & Reuter-Oppermann, M.	Veröffentlicht
P6		Eye-tracking the Lifeline: Emergency Dispatchers' Visual Attention in Time-Sensitive Calls	Manuskript einreichungsbereit („submission ready“)	Elsenbast, C.	Einreichungsbereit

1 Einleitung und Relevanz des Forschungsthemas

1.1 Heranführung

Behörden- und Organisationen mit Sicherheitsaufgaben (BOS), wie Rettungsdienst, Feuerwehr, Katastrophenschutz und Technisches Hilfswerk (THW), sind tragende Säulen der staatlichen Daseinsfürsorge. Hilfeersuchen, Notrufe und Einsätze der BOS werden in Leitstellen angenommen und koordiniert abgearbeitet. Dafür sind innerhalb der Leitstellen in der Regel mehrere BOS räumlich und funktionell zusammengefasst. Deutsche Leitstellen initiieren durchschnittlich alle drei Sekunden die Entsendung eines Notfallrettungsmittels (z. B. Rettungswagen) und alle sechs Sekunden die Alarmierung eines Fahrzeugs für einen Krankentransport (Baubin et al., 2020, S. 487).

Im Rahmen dieser Arbeit sind mit „Leitstellen“ die Lage-, Kontroll- und Koordinierungsorgane der BOS gemeint, sofern nicht explizit anders spezifiziert. In der Regel handelt es sich um sog. „Integrierte Leitstellen (ILS)“ für Brandschutz, Katastrophenschutz und Rettungsdienst, die ständig mit Personal besetzt sind und entsprechende Informations- und Kommunikationsmittel bereithalten, um Notrufe entgegenzunehmen, Ressourcen der Gefahrenabwehr zu koordinieren und um die Verbindung mit anderen BOS, Leitstellen oder relevanten Dritten, z. B. Krankenhäusern, herzustellen. Die Leitstellen sind ein elementarer Bestandteil der kritischen Infrastrukturen (KRITIS) sowie der Daseinsfürsorge und sind mit Aufgaben betraut, die in die Hoheit des Staates unter behördlicher Zuständigkeit der Bundesländer fallen (Trautmann et al., 2022, S. 2–3).

Die Leitstellenlandschaft in Deutschland als ist generell „heterogen“ zu beschreiben. Die Gebietsgröße des Zuständigkeitsbereichs einer Leitstelle ist beispielsweise in den deutschen Flächenstaaten sehr unterschiedlich, so beträgt die größte Gebietsgröße 5961,70 km² in Brandenburg und die kleinste 650,02 km² in Nordrhein-Westfalen. In einer Befragung der Leitstellenleitungen (n=55) konnten unterschiedlichste Konzepte zu Trägerschaft gezeigt werden (Trautmann et al., 2022, S. 35). Auch das Verfahren des Notrufdialogs unterscheidet sich zwischen den Leitstellen erheblich. So bestehen im Grundsatz drei Möglichkeiten der Notrufabfrage: 1. keine Vorgaben zur Strukturierung der Notrufabfrage, die Leitstellendisponent:innen fragen „frei“ und nach persönlich gesetzten „Verfahren“ ab; 2. strukturierte Notrufabfrage mit einer vorgegebenen Struktur, die jedoch durch die Disponent:innen selbstständig angepasst werden kann, ggf. softwaregestützt mit automatischer Dokumentation und optionaler automatischer Übernahme in das sog.

Einsatzleitsystem² (ELS); 3. standardisierte Notrufabfrage (SNA), bei der der genaue Wortlaut verpflichtend zu übernehmen ist, mit obligater Softwareunterstützung und Übernahme ins ELS. Aus der Unterscheidung zwischen strukturierter und standardisierter Notrufabfrage entstehen neben dem Erzielen einer personenunabhängigen Prozessqualität weitere Konsequenzen, wie Unterschiede in der Evaluierbarkeit und Rechtssicherheit, weshalb die SNA als Goldstandard zu betrachten ist (Mayr, 2020). Die SNA ist derzeit bundesweit nur lückenhaft umgesetzt und Maßnahmen zum Monitoren von qualitätskritischen Parametern, wie z.B. Notrufannahmezeit und Notrufbearbeitungszeit, scheinen weitgehend nicht erhoben zu werden, obwohl nicht von einer personellen Mehrbelastung durch Implementierung der SNA auszugehen ist (Luiz et al., 2019, S. 286–292).

Auch hinsichtlich interner Arbeitsprozesse unterscheiden sich die Leitstellen mitunter deutlich. So ist in der internen Ablauforganisation zwischen Calltaker-Dispatcher-Systemen (CDS) und Einsatzsachbearbeiter-Systemen (ESS) zu unterscheiden (Trautmann et al., 2022, S. 12). Calltaker nehmen Notrufe entgegen und geben bei Bedarf weitere Anweisungen und Hilfestellungen. Die Informationen aus dem Notrufgespräch werden digital im ELS erfasst. Auf dieser Informationsgrundlage alarmieren die Dispatcher das geeignete Rettungsmittel und koordinieren den weiteren Einsatz. Das ESS verzichtet auf diese Aufgabentrennung. Einsatzsachbearbeiter:innen sind also Calltaker und Dispatcher gleichzeitig. Zur Vereinfachung wird sich in der vorliegenden Arbeit auf die zusammenfassende Bezeichnung Leitstellendisponent:innen oder verkürzt Disponent:innen beschränkt, sofern die Unterscheidung zwischen CDS und ESS nicht im jeweiligen Kontext relevant sein sollte.

Außerdem bestehen weitere Unterschiede zwischen den Leitstellen, wie beispielsweise unterschiedliche Schichtlängen sowie interne Personalrotationen, bei denen die Disponent:innen noch andere Aufgaben in einer „leitstellenfreien“ Bereitschaftszeit übernehmen (z. B. die Besetzung von Sonderfahrzeugen).

1.2 Technik- und menschenzentrierte Perspektiven auf KI in den Leitstellen

Um die Leitstellendisponent:innen bei den herausfordernden Aufgaben in einer heterogenen Arbeitslandschaft bei den komplexen Aufgaben zu unterstützen und somit die Patient:innensicherheit und Versorgungsqualität zu steigern, stehen unterschiedliche technische Hilfsmittel zur Verfügung, wie computergestützte ELS und weitere Informations-

² Ein ELS ist eine spezialisierte Software zur Koordination und Verwaltung von Notrufen, Einsatzkräften und Ressourcen bei Behörden und Organisationen mit Sicherheitsaufgaben.

und Kommunikationstechnologien, die teilweise außerhalb der ELS zu bedienen sind, wie Echtzeitverkehrsinformationen oder Anwendungen der Wetterdienste. Die Leitstellendisponent:innen haben dabei die Aufgabe, Notrufe anzunehmen und vielfältige Informationen zu einer einsatztaktischen Entscheidung, wie der Alarmierung der richtigen Rettungsmittel, zusammenzuführen. Außerdem besteht einer der erschwerenden Faktoren der Leitstellenarbeit in der häufig mangelnden (technischen) Interoperabilität der Leitstellen untereinander und zu weiteren relevanten Organisationen (z. B. Polizei). Die Vernetzung ist also kaum über den eigenen Zuständigkeitsbereich hinweg gegeben.

Hinsichtlich der Anwendung von Verfahren der künstlichen Intelligenz (KI) in den Leitstellen bemerken Krafft et al. (2022, S. 109), dass es zwar international vielversprechende Anwendungen gebe, aber „Deutschland bei dieser Entwicklung bisher nicht bemerkenswert in Erscheinung getreten“ ist. Da KI-Verfahren in der Leitstelle jedoch ein hohes Potenzial zugesprochen wird (Dong et al., 2022; Emami & Javanmardi, 2023; Scholz et al., 2022), mag das erstaunen. Ein konkretes Beispiel für ebenjenes Potenzial wird in der Studie von Blomberg et al. (2019) dargestellt. Blomberg et al. zeigen auf, dass die Verwendung einer Spracherkennungssoftware bei Notrufen, die die Disponent:innen durch Hinweise unterstützt, jedoch die Entscheidungen bei den Disponent:innen belässt, zu einer früheren Erkennung des Kreislaufstillstandes geführt hat, im Vergleich zu Notrufen ohne KI-Anwendung. Neben der klinischen Entscheidungsunterstützung in weiteren Situationen führen Krafft et al. (2022, S. 111) auch das Potenzial der KI-Verfahren in Leitstellen zur Qualitätssicherung an. So kann die automatische Analyse der immensen und vielfältigen Leitstellendaten einen wesentlichen Beitrag zur Prozesssteuerung und -optimierung leisten, wie Krafft et al. ebenda betonen.

Auch Maletzki et al. (2023) gehen in ihrer Studie einerseits auf den aktuellen Stand von KI-Verfahren in den Leitstellen ein und beschreiben andererseits weitere mögliche Anwendungen. Die Autor:innen zeigen auf, dass per regelbasiertem Schließen der Bewusstseinszustand von Patient:innen mit der Atemfrequenz in Verbindung gebracht werden kann, um einen Herz-Kreislauf-Stillstand zu erkennen. In der Praxis können diese Modelle jedoch versagen, was in falschen Empfehlungen resultiert. Die Disponent:innen müssen dann ihre Erfahrungen aus anderen Fällen nutzen, um zu erkennen, wann einer regelbasierten Schlussfolgerung des Systems zu widersprechen ist. Weitere Unterstützung für Disponent:innen kann durch den Einsatz künstlicher neuronaler Netze gefunden werden, die jedoch in der Praxis zu einer erheblichen Anzahl von Fehlalarmen führen können (Blomberg et al., 2021; Byrsell et al., 2021). Maletzki et al. (2023) erwarten daher, dass künstliche

neuronale Netzwerke besonders hilfreich sind, um einen schnellen Hinweis auf eine spezifische Diagnose zu geben, die anschließend durch andere KI-Methoden weiter untersucht werden muss. Derzeit bestehen vor allem Lücken in möglichst umfassenden KI-Lösungen, die mehr als nur sehr spezifisch umschriebene Szenarien abbilden. Deshalb sind derzeit verschiedene Insellösungen im Bereich der KI miteinander zu verbinden, um verschiedene Szenarien und Zwecksetzungen abdecken zu können.

Betrachtet man die Anwendung von KI-Verfahren in den Leitstellen aus menschenzentrierter Sicht, ist zunächst hervorzuheben, dass die Evidenz für den Leitstellenbereich weitgehend fehlt. Eine menschenzentrierte Ausrichtung der Anwendungsforschung bedeutet in Anlehnung an Große (2020), dass Lösungen für Probleme entwickelt werden, die von den betroffenen Menschen auch gerne genutzt werden, weil sie als hilfreich empfunden werden. Menschenzentriert zu forschen, heißt in diesem Bezug,

„die Probleme, Anforderungen und Lebenssituationen der betroffenen Menschen zu kennen. Es bedeutet, die Sprache zu verwenden, die diese Menschen sprechen und Lösungen so zu gestalten, dass sie intuitiv zu bedienen sind. Es bedeutet, iterativ zu arbeiten, Prototypen immer wieder zu testen und Entwürfe zu verwerfen, um am Ende ein gutes Ergebnis zu haben.“
(Große, 2020)

Dass menschenzentrierte Aspekte der Mensch-KI-Interaktion in den Leitstellen in diesem Maße unterbeforscht sind, überrascht auch aus der Perspektive, dass die Disponent:innen regelmäßig weitreichende Entscheidungen treffen müssen, die Leben, Gesundheit, öffentliche Sicherheit oder hohe Sachgüter betreffen. Besondere Brisanz gewinnt diese menschenzentrierte Perspektive auch dadurch, dass Disponent:innen sichere und schnelle Entscheidungen regelmäßig unter Unsicherheit, Zeitdruck und Informationsdefiziten treffen müssen (siehe Kapitel 2).

1.3 Ziele und Forschungsfragen

Die vorliegende Arbeit hat das **Ziel**, menschenzentrierte Aspekte der Leitstellenarbeit inklusive deren Nahtstellen (z.B. zum Einsatzdienst von Feuerwehr und Rettungsdienst) zu beforschen, um eine Grundlage für die Entwicklung von Unterstützungssystemen (z. B. KI-Verfahren) zu legen, die möglichst exakt die Bedürfnisse der Leitstellendisponent:innen treffen.

Es werden im Verlauf der Arbeit drei **inhaltliche Probleme** adressiert: **Erstens** besteht zur Wahrnehmung von Komplexität und Variabilität der Arbeitsumgebung von in Deutschland tätigen Leitstellendisponent:innen keine Datenlage. Datengetriebene Aussagen zum Status

quo inklusive Real-World-Problems der Disponent:innen können so nicht zuverlässig getroffen werden, was auch die Entwicklung von unterstützenden Technologien nachteilig beeinflusst. Da Komplexität durch das Auslösen von Stress und Angst kognitive Prozesse behindern und somit adäquate Reaktionen verhindern kann (Dörner, 2008, S. 288), ist dieses Problem aufgrund der folgenreichen Entscheidungen der Disponent:innen hochrelevant. **Zweitens** ist auch zur Technikbereitschaft der Disponent:innen und zur Akzeptanz neuer Technologien keine wissenschaftliche Literatur zu finden. Diese Erkenntnisse sind jedoch von hohem Stellenwert, um zu verstehen, welche Faktoren die erfolgreiche Implementierung und Nutzung intelligenter Systeme beeinflussen. Im Bereich der Krankenpflege stellten Strutz et al. (2020), u. a. im Rückgriff auf Hobbs (2002), mangelnde Technikakzeptanz als Innovationshürde heraus und beschrieben, dass Pflegekräfte, die über eine positive Einstellung und Fähigkeiten im Umgang mit Informationstechnologien verfügen, auch eine größere Technikkompetenz aufweisen als andere Pflegekräfte. **Drittens** sind zur visuellen Aufmerksamkeitsverteilung und zur Interaktion der Disponent:innen mit dem ELS keine Daten verfügbar. Dies ist wiederum wichtig, um einschätzen zu können, welche visuellen Informationen essenziell für die Entscheidungsfindung sind und welche distrahierend wirken. Diese Erkenntnisse können außerdem Ansätze zu intelligenten und intuitiv zu verarbeitenden Visualisierungsstrategien aufzeigen. Die einzige verfügbare Arbeit zu den mental-informativischen Anforderungen im Leitstellensetting weist auf eine hohe Belastung der Disponent:innen hin, was den Stellenwert unterstützender Technologien untermauert (Herbig & Müller, 2016).

Der **methodologische Rahmen** der Studie basiert auf einem erweiterten Mixed-Methods-Design, das quantitative, qualitative und physiologische Erhebungsverfahren integriert. Das gewählte Design folgt der Komplementaritätsthese (Döring et al., 2015), welches von einer epistemologischen Abgrenzung zwischen quantitativem und qualitativem Forschungsparadigma ausgeht, dabei jedoch auf deren wechselseitige Gleichwertigkeit abstellt. Die inhärenten methodischen Limitationen der einzelnen Paradigmen werden durch die systematische Integration vollständiger qualitativer und quantitativer Teilstudien adressiert. Dieses Mixed-Methods-Research-Design wird durch die Implementierung physiologischer Messverfahren erweitert, um eine mehrdimensionale Erfassung des Forschungsgegenstandes zu ermöglichen.

Die dargestellten drei Problemstellungen führen zu den drei **Forschungsfragen** dieser kumulativen Dissertation:

1. Wie ist die Zusammensetzung der Zielgruppe im Bereich der Leitstellendisponent:innen und weiteren Anwender:innen im Schnittstellenbereich³ hinsichtlich soziodemografischer Merkmale und spezieller Merkmale, wie Team-Arbeit-Kontext und Technikbereitschaft, und welche Implikationen ergeben sich aus diesen Erkenntnissen für die Entwicklung und Evaluation intelligenter Systeme?
2. Wie lässt sich das Überzeugungssystem der Leitstellendisponent:innen und weiteren Anwender:innen im Schnittstellenbereich in Bezug auf KI-Implementierungen und vernetzte, intelligente Plattformen beschreiben?
3. Auf welche grafischen Elemente verteilt sich die Aufmerksamkeit der Leitstellendisponent:innen in der Phase von der Annahme eines Notrufs bis zum Ende der Alarmierung?

Im weiteren Verlauf dieser Arbeit wird eine Einordnung der wichtigsten Konstrukte vorgenommen, der methodologische Rahmen dargestellt, die Forschungsfragen begründet, die Publikationen und Ergebnisse zusammengefasst und zur Diskussion übergeleitet.

2 Komplexität in Leitstellen

Die Tätigkeit der Leitstellendisponent:innen ist durch Komplexität charakterisiert. Da dies eine grundlegende Annahme für die vorliegende Dissertation ist, wird das Thema in einem eigenständigen Kapitel elaboriert.

Ein komplexes System besteht im Rückgriff auf Dörner (2008, S. 285) aus vielen Variablen, die unüberschaubar miteinander verknüpft sind. Dies führt dazu, dass komplexe Systeme chaotisch erscheinen und „dass unter ähnlichen Verhältnissen oft nicht ähnliche Inputs zu dem gleichen Effekt führen“ (Dörner, 2008, S. 285), woraus ein hohes Maß an Intransparenz entsteht. Komplexität entsteht auch in Leitstellen durch eine eigendynamische Verknüpfung vielfältiger Faktoren, die vor dem Hintergrund limitierter Informationen und zeitkritischer Entscheidungsnotwendigkeiten zu verstehen sind. Dörner (2008, S. 286) weist auch auf die zeitliche Dimension von Komplexität hin, was vor allem im Leitstellenbereich durch zeitkritische Notfälle an Brisanz gewinnt:

„Man hat auch nicht genügend Zeit, um ausgiebig Erkundungen zu betreiben, denn die Realität wartet nicht, bis wir unsere Explorationen beendet haben, sondern entwickelt sich weiter.“

³ Als „weitere Anwender:innen im Schnittstellenbereich“ werden im Rahmen dieser Arbeit vor allem die Einsatzkräfte der BOS und Gesundheitseinrichtungen verstanden, die an der Schnittstelle Rettungsdienst/Feuerwehr/Krankenhaus und Leitstelle arbeiten. Aus dem Begriff „Schnittstelle“ wurde im Verlauf dieser Arbeit der Begriff „Nahtstelle“, um den gemeinsamen und interprofessionellen Charakter der Interaktion auch semantisch zu berücksichtigen und um nicht primär auf Unterschiede abzustellen.

Bei der konkreten Betrachtung von Leitstellen zeigen sich verschiedene komplexitätssteigernde Faktoren. Ein wesentlicher Aspekt ist dabei das Informationsgefälle zwischen Anrufer:in und Disponent:in: Die Hilfesuchenden müssen ihr Anliegen möglichst knapp und präzise schildern und ebenso prägnant auf Rückfragen des Leitstellenpersonals antworten. Ein ruhiges, rationales Verhalten steht jedoch oftmals im diametralen Kontrast zur Affektlage der Anrufer:innen, denen oftmals die kritische Situation und die eigene Hilflosigkeit wohlbewusst sind. Weiterhin ist es den Disponent:innen in der Regel nicht möglich, die Situation am Notfallort selbst in Augenschein zu nehmen. So entstehen eine erhebliche Intransparenz sowie ein Missverhältnis aus zur Verfügung stehenden und benötigten Informationen, um in kürzester Zeit eine irreversible einsatztaktische Entscheidung zu treffen, die bedeutende Folgen für Leben, Leib, Freiheit, Umwelt oder Sachgüter haben wird. Weiterhin ist in einer Vielzahl der Notrufe von einer dynamischen Situation auszugehen, beispielsweise kann sich der Gesundheitszustand von Patient:innen rapide verschlechtern oder eine Ausbreitung der Gefahr drohen (z.B. Brand, Gasaustritt etc.). Folglich müssen unter Zeitdruck unsichere Entscheidungen bei begrenzten materiellen und kognitiven Ressourcen, von den Leitstellendisponent:innen und Einsatzkräften getroffen werden.

Die beschriebene Komplexität macht Menschen Angst, wie auch Dörner ausführt (2008, S. 286), der diese Angst mit der Wahrnehmung von Unbestimmtheit und Hilflosigkeit verbindet, die zu entsprechenden Angstreaktionen führen können, wie dem imperativen Drang, sich der auslösenden Situation zu entziehen oder diese zu bekämpfen. Zudem muss noch ein Augenmerk auf sog. Sicherungsverhalten gelegt werden. Sicherungsverhalten ist die Tendenz, das mentale Bild der Situation (z. B. Lage am Einsatzort), als wie auch mentale Bilder der antizipierten Entwicklungsmöglichkeiten der Situation möglichst genau zu berechnen (Dörner, 2008, S. 286). Dies kann jedoch, ähnlich wie Flucht- und Aggressionstendenzen auch, höchst problematisch sein, da andere kognitive Ressourcen, die für die Entscheidungsfindung benötigt werden, aufgrund der limitierten Denk- und Aufmerksamkeitsressourcen blockiert werden. Die überlagerten kognitiven Ressourcen können aus den Fähigkeiten zum angemessenen Strategieeinsatz, der Anpassung von automatischen Reaktionen, der Feedbackverarbeitung und der Sensibilität für Belohnung und Bestrafung bestehen (Starcke & Brand, 2012, S. 1241). Stress (beispielsweise als Folge von Angst) an sich führt bereits über eine Minderung des sog. Auflösungsgrades zu kognitiven Einschränkungen. Der Auflösungsgrad bezeichnet das Ausmaß der Ausarbeitung kognitiver Prozesse, einschließlich

der Genauigkeit der Wahrnehmung, der Tiefe und Breite der Erinnerung sowie der Konkretheit der Erinnerung (Dörner, 2008, S. 288). Er umfasst auch das Maß, in dem bei der Planung von Handlungen Fern- und Nebenwirkungen berücksichtigt werden, sowie die Präzision, mit der überprüft wird, ob die Bedingungen für die erfolgreiche Durchführung einer Aktion gegeben sind. Diese Annahmen über die Auswirkungen von Stress werden auch von Phillips-Wren und Adya (2020) gestützt, die Informationsüberlastung, Zeitdruck, Komplexität und Unsicherheit als Hauptprobleme bei Entscheidungsfindungen identifizierten. Alle vier Probleme können sich bei der Leitstellenarbeit zeitgleich manifestieren. Wie Hofinger (2003, S. 17) ergänzt, antworten Menschen zur Vereinfachung und zum Selbstschutz in diesen Hochstresssituationen mit einer Vielzahl von Strategien, die das eigentliche inhaltliche Handlungsziel implizit unterpriorisieren. Hofinger (2017, S. 104–105) schreibt dazu:

„Einsätze im Rettungsdienst können vorausgeplant werden. Die nötigen Entscheidungen werden als Handlungsanweisungen festgehalten. Je nach Anwendung können dies Standardprozeduren sein [...], Algorithmen oder Leitlinien sowie Handbücher etc. All diesen Formen ist gemein, dass sie die Personen, [...], durch inhaltliche Vorgaben beim (Aus-)Suchen möglicher Herangehensweisen entlasten.“

Die Prämisse der Vorausplanbarkeit rettungsdienstlicher Einsätze mit konsekutiver Entlastung der Einsatzkräfte durch algorithmisierte Vorgaben oder Handbücher kann jedoch nicht gänzlich unkritisiert bleiben, denn sie verkennt das Ausmaß der kognitiven Prozesse (beispielsweise durch Sicherungsverhalten, Herbeiführen einer initialen Arbeitsdiagnose oder Reflexion eines angemessenen Strategieeinsatzes), die benötigt werden, um den zutreffenden Algorithmus auszuwählen. Dies ist vor allem darin begründet, dass die Disponent:innen in kürzester Zeit die Informationen der Notrufer:innen verstehen und priorisieren müssen, während parallel die Eingabe in die Notrufannahmemaske des ELS gemacht wird, bevor aus einer Vielzahl von Algorithmen der am besten passende teilweise manuell durch die Disponent:innen ausgesucht werden muss. Dazu ist durch die Disponent:innen auf der Grundlage limitierter Informationen eine möglichst präzise Verdachtsdiagnose zu erstellen oder zumindest das Hauptsymptom zu erkennen und priorisiert auszuwählen. Erst von den weiteren Prozessen der Entscheidungsunterstützung, wie einem entsprechenden Rettungsmittelvorschlag, der neben der Art des Rettungsmittels auch geografische Informationen beinhaltet und durch die Disponent:innen zu prüfen ist, ist eine kognitive Entlastung anzunehmen. Hier entsteht bei der Allokation von Aufmerksamkeit bereits ein Zielkonflikt, der kognitive Kosten (mentale Arbeitsbelastung) mit sich bringt. Hofinger (2003) verweist auf die begrenzte Kapazität kognitiver Prozesse bei der sprachlichen Verarbeitung innerhalb kurzer Zeitintervalle. Diese Limitation ist für die parallele Durchführung von Notrufdialog und Selektion relevanter Handlungsanweisungen von besonderer Bedeutung.

Folglich erfordert ein KI-basiertes Entscheidungsunterstützungssystem eine automatisierte, kontextsensitive Informationsbereitstellung, die relevante Daten zeitgerecht präsentiert und simultan irrelevante Informationen sowie Störfaktoren filtert. Kontextsensitivität beschreibt hier den zu entwickelnden Mechanismus des ELS, der auf der grafischen Bedienoberfläche nur die diejenigen Informationen präsentiert, die auch zur Bewältigung der jeweiligen, notruf- bzw. einsatzspezifischen Situation erforderlich sind.

3 Methodologischer Rahmen

3.1 Verortung des Forschungshintergrundes mit methodologischen Erwägungen

Das vom BMWK öffentlich geförderte Projekt „SPELL“ (Förderkennzeichen 01MK21005B) ist der praktische Forschungshintergrund dieser Arbeit. SPELL steht für Semantische Plattform zur intelligenten Entscheidungs- und Einsatzunterstützung in Leitstellen und Lagezentren. In SPELL wird durch ein gemeinsames digitales Ökosystem⁴, auf das unterschiedliche Leitstellen durch eine niederschwellige Schnittstelle zugreifen können, Probleme der mangelnden Interoperabilität gelöst. Auf diesem digitalen Ökosystem sind beispielsweise Dienste, die auf Verfahren der KI basieren, gelagert und somit für die Nutzer:innen unkompliziert zugänglich. Ein Beispiel für die Arbeit von SPELL ist die KI-Anwendung „KINAS“, die einen sprachbasierten Simultanübersetzer für Notrufe darstellt. Das fremdsprachliche Gesagte der Notrufer:innen wird direkt übersetzt und als Sprachausgabe ausgegeben, sowie vice versa, also die Fragen der Leitstellendisponent:innen in die Fremdsprache übersetzt. Dabei sind diese Systeme weder als vollautomatische Substitutionsversuche der Disponent:innen zu verstehen, noch sollen sie ohne Eingriffsmöglichkeit eigenständig strategische Entscheidungen treffen. Vielmehr sollen sie dazu führen, dass einsatztaktische relevante und erfolgskritische Aufgaben nur als Vorschlag umgesetzt werden, wenn der Mensch ihn annimmt oder vor der Automatisierung der Entscheidung ein Veto eingelegt werden kann. Angelehnt an Parasuramans (2000) zehn Stufen der Automatisierung, wären diese Dienste auf Stufe 5 bzw. 6 (siehe Abbildung 1) zu sehen.

Um solche KI-Mehrwertdienste zielspezifisch und menschengerecht zu entwickeln, müssen die Bedürfnisse, Anforderungen und Haltungen der Leitstellendisponent:innen und weiterer Akteur:innen, wie z. B. Notfallsanitäter:innen und Notärzt:innen, gegenüber der Leitstellenarbeit und Technologien auf KI-Basis erhoben und analysiert werden, um die KI-

⁴ In einem digitalen Ökosystem sind verschiedene digitale Technologien, Plattformen und Dienste miteinander verbunden, um durch die Verknüpfung der einzelnen Systemteile synergistische Effekte für die Nutzer:innen zu erzielen. Digitale Ökosysteme sind als „Platform as a Service“ zu verstehen.

Dienste passgenau entwickeln, integrieren und evaluieren zu können. Diese Aspekte werden in der vorliegenden Dissertation als zentrale Themen adressiert.

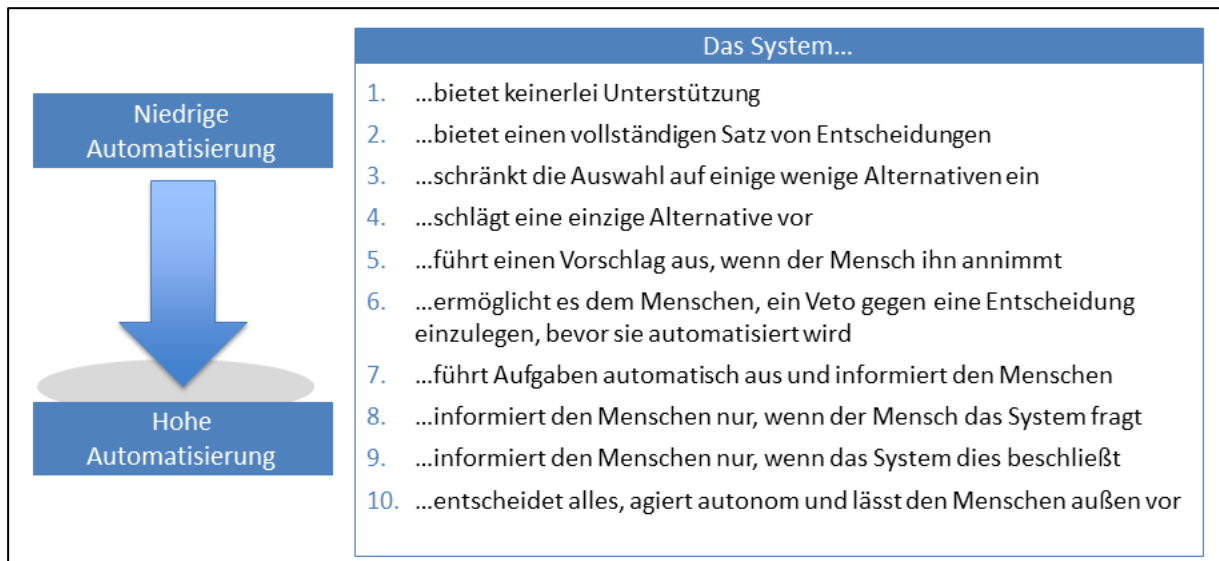


Abbildung 1: Stufen der Automatisierung angelehnt an Parasuraman (2020)

3.2 Zentrale Begriffe und Konzepte

3.2.1 Technikbereitschaft

Als vorwegnehmende Erklärung kann zur Orientierung festgehalten werden, dass Technikbereitschaft nach Neyer et al. (2012) ein Konzept ist, das die Bereitschaft zum Umgang mit Technik anhand von drei Facetten erfasst: Technikakzeptanz, Technikkompetenzüberzeugung und Technikkontrollüberzeugungen. Betrachtet man ein entwickeltes Werkzeug, soll dieses nach Fertigstellung auch von der Zielgruppe verwendet werden. Im Optimalfall herrscht eine hohe intrinsische Motivation zur Nutzung, die Adressat:innen sind mit dem Werkzeug zufrieden, empfinden es als Arbeitserleichterung und eine sog. Nachkaufreue der Verantwortlichen ist nicht vorhanden. Somit ist bereits im Entwicklungsprozess zu hinterfragen, welche Eigenschaften und Funktionalitäten ein Werkzeug haben muss, um tatsächlich vorhandene Probleme bei den Adressat:innen zu lösen und auf deren individuelle Bedürfnisse hin zu entwickeln. Versteht man Verfahren der KI, genauer Entscheidungsunterstützungssysteme, als Werkzeuge, dann treffen diese Überlegungen ebenso zu. Der folgende Abschnitt beschäftigt sich mit der Perspektive der menschenzentrierten Faktoren, die eine erfolgreiche Werkzeugnutzung vorhersagen.

Das Technology Acceptance Model (TAM) (Davis, 1985) ist ein etabliertes und viel genutztes Instrument zur Untersuchung der Technologienutzung (Schepers & Wetzels, 2007). Das TAM basiert auf der Annahme, dass die Einstellung gegenüber einer Technologie (Attitude toward Using) durch zwei Kernvariablen determiniert wird: die wahrgenommene Nützlichkeit

(Perceived Usefulness) sowie die wahrgenommene Benutzerfreundlichkeit (Perceived Ease of Use). Erstere beschreibt die Erwartung, dass die Technologienutzung in einem Benefit für die jeweilige Person aufgeht. In dieser Annahme spiegelt sich auch die Arbeit einiger Motivationstheorien, wie beispielsweise der VIE-Theorie (Nerdinger et al., 2018, S. 476–478). Die Theorie beschreibt den Zusammenhang zwischen der Erwartung eines Nutzens und der Motivation zur Nutzung einer Technologie bzw. allgemeiner formuliert zur Initiierung einer Handlung. Kurzum wird demnach Leistung von Menschen nur dann als erstrebenswert wahrgenommen, wenn es wahrscheinlich ist, damit ein erwünschtes Ziel zu erreichen. Zurückkehrend zum TAM wird unter Perceived Ease of Use das Ausmaß beschrieben, in dem Menschen den Aufwand zum Erlernen der Technologienutzung bewerten. Die Intention to Use, also letztendlich die verhaltensbedingte Nutzungsabsicht, hängt wiederum von den drei vorgestellten Konstrukten ab.

Neyer et al. (2012) kritisieren den Ansatz des TAM bezüglich seiner fehlenden Rücksichtnahme auf Faktoren der menschlichen Persönlichkeit. Kompetenz- und Kontrollüberzeugungen werden außer Acht gelassen und Persönlichkeitseinflüsse spielen kaum eine Rolle. So finden sich im Modell von Neyer et al. (2012) die Facetten Technikakzeptanz, Technikkompetenz- und Technikkontrollüberzeugungen. Neyer et al. (2012) beschreiben deutlich, dass individuelle Unterschiede der Technikbereitschaft nicht nur substanzial mit Persönlichkeitsmerkmalen kovariieren, sondern auch Verhalten, wie Techniknutzung und Technikbiografie, vorhersagen können (Neyer et al., 2012). Als Beispiel für die grundlegenden Persönlichkeitsmerkmale nennen sie emotionale Stabilität und Offenheit für Erfahrungen. Hingegen würde der Schulabschluss zwar substanzial mit generalisierter Kompetenz- und Kontrollüberzeugung kovariieren, Intelligenzmaße würden die Kovarianz jedoch nur äußerst schwach aufgreifen. Demzufolge wäre Technikbereitschaft zwar qualifikationsabhängig, jedoch unabhängig von der Intelligenz. Zur Messung der Technikbereitschaft entwickelten und validierten Neyer et al. (2012) einen Fragebogen, der vier Fragen pro Konstrukt beinhaltet und die Auswertung der drei Facetten erlaubt, aber auch Technikbereitschaft als ungewichtete Summierung der drei Facetten berücksichtigt. Wichtig zu erwähnen ist, dass sich das Konzept von Neyer et al. (2012) auf ein breites Spektrum von Technologien und technischen Geräten bezieht, wie beispielsweise Unterhaltungs- und Kommunikationstechnik sowie Haushaltsgeräte, jedoch nicht auf intelligente Systeme, wie sie in der vorliegenden Arbeit adressiert werden.

3.2.2 Team-Arbeit-Kontext

Bereits die Forschung mit der „Kurzskala Technikbereitschaft“ rückt menschenzentrierte Aspekte der Leitstellen und des Rettungsdienstes ins Zentrum. Die Wahrnehmung von in Deutschland tätigen Leitstellendisponent:innen, Rettungsfachpersonal und weiteren Akteur:innen im Leitstellenbereich bezüglich tätigkeitsimmanenter Arbeitskontexte ist jedoch bislang wenig beforscht. Aus diesem Grund fehlen für technologiegetriebene Entwicklungen belastbare empirische Ansätze der Pain-Point-Analyse⁵ und der datenbasierten Entwicklung von Gains weitgehend. Diese Lücken sollen durch die Anwendung des Team-Arbeit-Kontext-Inventars (TAKAI) besonders für die Leitstellendisponent:innen geschlossen werden.

Der Ursprung des TAKAI liegt in der Forschung von Hagemann et al. (2011), die den Arbeitskontext von unterschiedlichen High-Responsibility-Teams (HRTs) vor dem Hintergrund des Kompetenzerwerbs beforschten. Hagemann et al. heben hervor, dass klassische Ansätze zur Weiterentwicklung von Gruppenprozessen nicht zu den Bedürfnissen der HRTs passen. Beispielsweise können HRTs, die als Ad-hoc-Teams mit spezifischen Leistungsindikatoren zu verstehen sind, keine klassischen Teamentwicklungsphasen (z.B. nach Tuckman, 1965) durchlaufen. Als Ergebnis der Fragebogenstudie wurden von Hagemann et al. (2011) 21 Analyseaspekte identifiziert und vier Kategorien zugeordnet, die in einem Einschließungsverhältnis stehen. Die Studie hat weiterhin gezeigt, dass die Anforderungen an verschiedene HRTs unterschiedlich sind, was die Frage nach spezifischen Trainingsangeboten aufwirft. So kann beispielsweise identifiziert werden, welche spezifischen Aspekte der Kommunikation und Entscheidungsfindung trainiert werden müssen und an welchen Stellen im Teamarbeitsprozess Störungen wahrscheinlich sind.

Das Schließen der Lücken hinsichtlich der menschenzentrierten Perspektive ist essenziell, um eine Entwicklung von Produkten, hier von KI-basierten Entscheidungsunterstützungssystemen, menschenzentriert, auf Grundlage der Pains und Gains von Disponent:innen, zu erlauben (siehe Kapitel 1.2). Die Anwendung des TAKAI schließt also weitere Lücken und eine tiefere Exploration der Arbeitskontexte von Disponent:innen wird ermöglicht.

3.2.3 Überzeugungssysteme

In Forschungsfrage 2 wird das Überzeugungssystem von Leitstellendisponent:innen und weiteren Anwender:innen hinsichtlich KI in Leitstellen beforscht.

⁵ Im Value Proposition Canvas werden mit Pain-Points, also Schmerzpunkten, Problembereiche eines Produkts definiert. Folglich können Gains (auch Gain-Points) als meist positiv konnotierte Effekte verstanden werden.

Ein **Überzeugungssystem** ist die Gesamtheit der persönlichen Auffassungen und Einstellungen, die über die Zeit hinweg relativ stabil und unabhängig von bestimmten Situationen bleiben (Heinecke-Müller, 2021). Diese Systeme beziehen sich meist auf spezifische Sachverhalte und beeinflussen die Wahrnehmung der Umwelt, der eigenen Person und das Zusammenspiel beider. Überzeugungssysteme bestehen aus kognitiven und affektiven Komponenten und können umfassende interne Repräsentationen von Kausalzusammenhängen bilden, wie Heinecke-Müller (2021) beschreibt. Sie sind nicht nur in der Sozialpsychologie, z. B. als Ursache von Vorurteilen, ein wichtiger Forschungsgegenstand, sondern spielen auch in vielen anderen psychologischen Teildisziplinen eine Rolle. Sie dienen als mögliche Operationalisierung für Subjektivität, wie im Fall der Attribution. Usó-Doménech und Nescolarde-Selva (2016, S. 117–148) erklären Überzeugungssysteme auch als die Geschichten, die wir uns selbst erzählen, um unser persönliches Verständnis von Realität zu definieren. Darüber hinaus zeigt Festingers (1957) Theorie der kognitiven Dissonanz, wie konfliktreiche Überzeugungen zu psychischem Unbehagen führen können, was Änderungen in Einstellungen oder Verhaltensweisen zur Wiederherstellung des Gleichgewichts anregt.

3.2.4 Visuelle Aufmerksamkeit

Selbst ein Gehirn von der Größe des menschlichen ist nicht ausreichend, um eine komplette Analyse der visuellen Informationsflut zu bewältigen, meinen Treue et al. (2007, S. 55). Davon ausgehend, dass Aufmerksamkeit zwei wichtige Funktionen hat, die auch für die Disponent:innen hochrelevant sind, nämlich die perzeptuelle oder gedächtnisbasierte Selektion bestimmter Inhalte oder Informationen sowie die darauf folgende Handlungsvermittlung (Allport, 1987), wird deshalb im Folgenden ein Überblick über ebenjene Selektionsvorgänge der visuellen Aufmerksamkeit gegeben. Dazu werden die beiden Hauptvertreter der Theorien visueller Aufmerksamkeitsverarbeitung vorgestellt.

Ortsbasierte Theorien der visuellen Aufmerksamkeit basieren auf dem Flankierreiz-Paradigma von Eriksen und Eriksen (1974) sowie dem Spatial-Cueing Paradigma von Posner (1980) und gehen davon aus, dass alle visuellen Stimuli eines bestimmten Bereichs auch ohne Blickfixation selektiert werden und dass sich Reize außerhalb dieses priorisierten Areals nicht oder nur eingeschränkt verarbeiten lassen (Krummenacher & Müller, o. D.; Müller & Krummenacher, 2012, S. 310–311). Ortsbezogenen Theorien zufolge funktioniert visuelle Aufmerksamkeit wie ein Lichtkegel, der es erlaubt ein bestimmtes Areal zu betrachten und damit der Informationsverarbeitung an diesem Ort Vorrang verschafft (Müller & Krummenacher, 2012, S. 312). Die Aufmerksamkeitslenkung kann reizgesteuert (exogen) oder

willentlich (endogen) geschehen. Der exogene Mechanismus besitzt das Potenzial, den endogenen Mechanismus zu dominieren, wodurch beispielsweise ein markanter exogener Reiz eine willentlich vorbereitete Handlungssequenz zu unterbrechen vermag (Müller & Rabbitt, 1989). Für die Leitstellenarbeit bedeutet dies, dass exogene Hinweisreize (z. B. Signallichtquellen) sparsam und sinnvoll einzusetzen sind, um ungewünschte Distractionen zu vermeiden. Außerdem muss eine Anordnung und visuelle Aufbereitung von Daten so erfolgen, dass die visuelle Informationsselektion erleichtert wird.

Im Rahmen **objektbasierter Theorien** wird postuliert, dass die visuelle Selektion nicht primär auf räumliche Positionen im visuellen Feld, sondern vielmehr auf die dort lokalisierten Objekte ausgerichtet ist. Kramer et al. (1997) zeigt auf, dass die initiale Objektselektion auf einer Repräsentation basiert, die sich als Struktur gruppierter Elemente zeigt. Diese Erkenntnis führt zu der Schlussfolgerung, dass die objektbasierte visuelle Selektion in ihrer Grundstruktur dennoch maßgeblich ortsbezogen operiert. Dies untermauert die These eines „Primats der ortsbezogenen Aufmerksamkeit“ (Müller & Krummenacher, 2012, S. 313).

3.3 Zentrale Methoden und methodologische Erwägungen

3.3.1 Fragebogenforschung und Psychometrie

Die Fragestellungen der Forschungsfrage 1 werden durch Fragebögen adressiert. Im Fokus stehen also insbesondere die soziodemografischen Merkmale, die Technikbereitschaft und der Team-Arbeit-Kontext der Zielgruppe sowie die Implikation der Erkenntnisse für die Entwicklung und Evaluation intelligenter Systeme.

Items und Testbatterien, die mehr als soziodemografische Parameter reflektieren, sind im Rahmen dieser Arbeit als psychometrische Instrumente zu verstehen. Diese lassen sich von Ad-hoc-Verfahren durch die entsprechende testtheoretische Absicherung unterscheiden (Stieglitz, 2014, S. 59). Von psychometrischen Verfahren wird dann gesprochen, wenn sie basierend auf der klassischen oder probabilistischen Testtheorie entwickelt wurden und entlang psychometrischer Gütekriterien validiert wurden. Im Rahmen dieser Arbeit wurden zwei psychometrische Instrumente eingesetzt, nämlich das Team-Arbeit-Kontext-Inventar (TAKAI) (Hagemann et al., 2011) und die Kurzskala Technikbereitschaft (Neyer et al., 2012). Letztere wurde durch den Autor weiterentwickelt und validiert. Die Validierung erfolgte beim TAKAI hauptsächlich durch die Berechnung der Intraklassenkorrelation (Koo & Li, 2016; Shrout & Fleiss, 1979) und die Multitrait-Multimethod-Technik (Campbell & Fiske, 1959), um einen Vergleich zur Arbeit von Hagemann et al. (2011) herstellen zu können. Die weiterentwickelte

Skala Technikbereitschaft wurde anhand der Konventionen von Hu und Bentler (1999) mit der konfirmatorischen Faktorenanalyse validiert.

Die Beantwortung der angesprochenen Forschungsfragen ist als eingebettet in den Prozess von Döring et al. (2015, S. 24–25) zu verstehen. Die Erstellung des soziodemografischen Fragebogenanteils beinhaltet die Darstellung fehlhandlungssicherer Items, z.B. Drop-down-Menü statt Freitexteingabe bei „Alter“.

3.3.2 Interviews und Grounded-Theory-Methodology

Zur Beantwortung der Forschungsfrage 2 wird ein Methodenmix vorgenommen und semistrukturierte Leitfrageninterviews per Expert:innen-Interviews sowie Fokusgruppeninterviews durchgeführt und nachfolgend mittels Gounded-Theory-Methodology (GTM) ausgewertet.

Das Expert:innen-Interview stellt eine spezifische Ausprägung des leitfadengestützten Interviews dar, bei welcher die Befragten in ihrer Funktion als Fachexpert:innen beteiligt werden, mit dem Ziel, deren domänenspezifisches Wissen zu erschließen und zu explizieren (Döring et al., 2015, S. 376). Die Interviewvariante richtet sich in pragmatischer Weise nach dem Gegenüber und den einzigartigen Charakteristika der jeweiligen Situation, welche wiederum von den Teilnehmenden, deren Erwartungen und Fähigkeiten sowie den theoretischen Vermittlungsannahmen der Forscher:innen abhängen, in Bezug darauf, was die Expert:innen zum Interview beisteuern können (Kassner & Wassermann, 2002, S. 106).

Im Rahmen dieser Arbeit sind Expert:innen Führungskräfte, wie Leiter von Leitstellen und Behörden. Leitstellendisponent:innen ohne Führungsaufgaben und weitere Positionen oder sonstige soziale Rahmungen stellen Spezialist:innen dar. Deshalb handelt es sich bei dieser Arbeit um Expert:innen- und Spezialist:innen Interviews. Für eine vertiefte Diskussion wird auf Anlage 3 verwiesen.

Der Wechsel zwischen Expert:innen- und Spezialist:innen-Interviews ist deshalb angezeigt, weil so auf den jeweiligen Stand der Theorieentwicklung mit dem passenden Interviewmodus reagiert werden kann.

Außerdem werden Fokusgruppen-Interviews durchgeführt, die als semi-strukturierte Varianten der Gruppendiskussion zu einem eingegrenzten Thema, hier intelligente Verfahren in Leitstellen, zu verstehen sind und bei denen eine hohe Teilnehmer:innen-Interaktion sowie eine hohe Selbstläufigkeit angestrebt werden (Döring et al., 2015, S. 380).

Mit **Grounded Theory** werden sowohl „eine Theorie als Ergebnis eines datennah erfolgenden Forschungsprozesses als auch dessen methodisches Prozedere selbst“ verknüpft (Pentzold et al., 2018, S. 2). Angesichts des Mangels an systematisierten und zuverlässigen Daten zum Zeitpunkt dieser Arbeit hinsichtlich Überzeugungssystemen der Leitstellendisponent:innen und weiterer Anwender:innen im Leitstellenkontext erscheint die GTM besonders geeignet, um eine Theorie über Verhaltensmuster und Einstellungen zu initiieren und zu entwickeln (Jones, 2009; McCallin, 2003). Die GTM kann auch verwendet werden, um Ängste und Sorgen zu erforschen und wie diese gelöst werden können (Glaser & Strauss, 2017; Pergert, 2009). Als interpretative Forschung wird die GTM oft von Sozialwissenschaftlern genutzt, um soziale Prozesse und Verhaltensweisen zu studieren (McCallin, 2003, S. 203). Interpretativ bedeutet jedoch nicht willkürlich, subjektiv und zufällig, sondern bezieht sich, im Hinblick auf die Qualitätskriterien dieser Arbeit, vielmehr darauf, dass Subjektivität im Zentrum steht. Die Theorie ist in den Daten selbst verankert (*grounded*), anstatt logisch oder hypothetisch abgeleitet zu werden (Jones, 2009, S. 30). Die GTM kann alle Arten von Daten im Forschungsprozess einbeziehen (Glaser & Strauss, 2017; Xie, 2009). Ein weiterer Vorteil des GTM-Ansatzes ist die Flexibilität in Bezug auf Datenerhebung und -analyse sowie die Fokussierung auf flexible und iterative, aber begründete Datenerhebung und -analyse (Glaser & Strauss, 2017; Jones, 2009). Ziel ist hier einerseits, ein Bild des Status quo zu generieren, das die Sicht der Expert:innen und Fachleute widerspiegelt, aber andererseits auch, eine realistische Vorstellung von der Zukunft zu extrahieren. Aufgrund des iterativen Prozesses der Datenerhebung, Kodierung und Analyse ist zu erwarten, dass immer wieder neue Fragen aus den Antworten entstehen, aber auch Erkenntnisse gewonnen werden, die die Ziele ermöglichen. Dieser iterative Prozess ist zu wiederholen, bis eine theoretische Sättigung erreicht wurde (Xie, 2009, S. 45). Somit hängt die Anzahl der Interviews (Stichprobengröße) vom Grad der theoretischen Sättigung ab.

Im Gegensatz zu quantitativer Forschung gibt es keine einheitlichen Qualitätskriterien für qualitative Ansätze (Yadav, 2022, S. 679), und die unkritische Übernahme klassischer quantitativer Kriterien (Objektivität, Reliabilität, Validität) wurde in der Vergangenheit kritisiert und abgelehnt (Mays & Pope, 2000, S. 50–51). Aus diesen Gründen orientiert sich diese Arbeit an Steinkes (2019, S. 319-331) sieben Hauptqualitätskriterien für qualitative Forschung, die über die Anforderungen der in der qualitativen Forschung weit verbreiteten

Kriterien von Lincoln und Guba (1985, S. 289–331) hinausgehen. Tabelle 2 gibt eine verkürzte Übersicht.

Tabelle 2: Gütekriterien qualitativer Forschung (nach Steinke, 2019, S. 319-331)

Kernkriterien	Hauptfragen
Intersubjektive Nachvollziehbarkeit	Inwieweit können Dritte den gesamten Forschungsprozess der Studie anhand der Studiendokumentation nachvollziehen und bewerten?
Indikation	Wie gut sind die einzelnen methodischen Entscheidungen im qualitativen Forschungsprozess der Studie begründet?
Empirische Verankerung	Wie gut sind die gebildeten Hypothesen und Theorien auf Basis der empirischen Daten gestützt?
Limitationen	Wie präzise wird angegeben, unter welchen weiteren Bedingungen die Forschungsergebnisse generalisiert werden können oder inwieweit die Generalisierbarkeit begrenzt ist?
Reflektierte Subjektivität	Wie umfassend und überzeugend reflektieren die Forschenden ihre eigenen subjektiven Positionen und Rollen im Verhältnis zum Phänomen und den untersuchten Personen?
Kohärenz	Wie kohärent und konsistent ist die Theorie oder die Interpretationen auf Basis der Daten?
Relevanz	Wie bedeutend ist der Beitrag einer qualitativen Grundlagenforschungsstudie für den wissenschaftlichen Fortschritt in Bezug auf Objektbeschreibung und Theoriebildung?

Die Analyse erfolgt mit MAXQDA2020 Analytics Pro [20.4.2, Verbi GmbH]. Der Verlauf wurde per Logbuch dokumentiert. Die Fragen der Interviews und die aggregierten Transkripte finden sich im hier zitierten Forschungsdatenrepositorium (Eisenbast, 2024).

3.3.3 Eyetracking

In Forschungsfrage 3 wird die visuelle Aufmerksamkeitsverteilung (hier: die Aufmerksamkeitsverteilung in Bezug auf die Monitore am Arbeitsplatz) der Leitstellendisponent:innen in der Phase von der Annahme eines Notrufs bis zum Ende der Alarmierung beforscht. Dazu werden Eyetracker eingesetzt. Eyetracker haben sich zu vielseitigen Geräten entwickelt, die zur Untersuchung verschiedenster Themen eingesetzt werden, wie Fahrzeugführer-Schnittstellen, Mensch-Computer-Interaktionen, Gaming und Softwareentwicklung (Sharafi et al., 2020, S. 1–2). Eyetracker ermöglichen es, signifikante und umfangreiche Daten über die Art und Weise aufzuzeichnen, wie Versuchspersonen mit visuellen Informationen interagieren, einschließlich Lesemustern (Rayner, 1998), visueller Hinweise während der Suche (Crosby et al., 2002) und Interaktionen und Engagement während mündlicher Gespräche (Bednarik et al., 2012). Sie werden auch in der Softwaretechnikforschung verwendet, um verschiedene Aufgaben zu untersuchen, wie z. B. das Lesen und Weiterverarbeiten von Quellcode, das Verständnis von Softwareartefakten und die Software-Traceability (Sharafi et al., 2020, S. 1–2). Einige Studien kombinierten

neurobildgebende und biometrische Techniken mit Eyetracking, um die Aufgabenschwierigkeit und kognitive Belastung zu messen (Fakhoury et al., 2018; Fritz et al., 2014; Lee et al., 2018).

Zum grundlegenden Verständnis der Technologie ist zu erwähnen, dass Eyetracking die Erfassung der visuellen Aufmerksamkeit der Versuchspersonen durch Aufzeichnung von Blickdaten beschreibt (Rayner, 1978 nach Sharafi et al., 2020, S. 3). Visuelle Aufmerksamkeit löst die kognitiven Prozesse aus, die für das Verständnis und die Problemlösung erforderlich sind, während kognitive Prozesse die visuelle Aufmerksamkeit auf spezifische Orte lenken. Daher ist Eyetracking nützlich, um die kognitiven Prozesse und die mentale Arbeitsbelastung der Versuchspersonen bei der Ausführung von Software- bzw. Softwareentwicklungsaufgaben zu untersuchen (Duchowski, 2003, S. 171–192). Im weiteren Sinne ist auch die Arbeit von Disponent:innen durch die starke Bindung an das ELS als Softwareaufgabe zu verstehen. Ein visueller Reiz ist jedes Objekt (beispielsweise eine Eingabemaske im ELS), das zur Ausführung einer Aufgabe notwendig ist und dessen visuelle Wahrnehmung die kognitiven Prozesse der Versuchsperson auslöst und letztendlich zu bestimmten Handlungen führt, z. B. dem Öffnen eines neuen Fensters in der Software. Dabei werden Eyetracking-Daten in Bezug auf bestimmte Bereiche eines Stimulus untersucht. Diese Interessengebiete werden als Areas of Interest (AOIs) bezeichnet und sind im Fall dieser Arbeit die verschiedenen Monitore der Leitstelle (Abbildung 2, rechts).

Bei der Umsetzung ist das spezielle Setting der Leitstellen zu berücksichtigen, beispielsweise hinsichtlich der Umgebungslichtbedingungen und der Lichtreflexionen der Monitore. Meist sind mehrere Bildschirme nebeneinandergeschaltet, teilweise auch zweireihig übereinander, teilweise mit weiteren großen Monitoren an den Wänden. Das Tracking muss also auch in Situationen robust abgebildet werden, in denen vielfältige Kopf- und Körperbewegungen zu verzeichnen sind. Durch die Monitorkonfigurationen beschränkt sich das Eyetracking auf mobile Systeme, also „Eyetracking-Brillen“, in diesem Fall das Produkt „Core“ der Firma Pupil Labs GmbH (o. D.) (Abbildung 2).

Das im Rahmen dieser Arbeit angewendete Set-up beinhaltet binokulares Eyetracking, es werden also beide Augen getrackt, in diesem Fall durch die robuste Dark Pupil Technique (DPT) (Min-Allah et al., 2021). DPT bedeutet, dass die Erfassung und Verarbeitung der Pupillen mit zwei Infrarotkameras erfolgen, die weiter entfernt von der optischen Achse der Kamera angeordnet sind (Tobii, 2023). Das Licht tritt in das Auge ein, wird von der Netzhaut reflektiert

und verlässt dann das Auge abseits der Achse und weg von der Kamera. Dies führt dazu, dass die Pupille im Bild dunkler erscheint als der Rest des Auges. Die Algorithmen zur Blickverfolgung identifizieren die Pupille, indem sie im Augenbild nach einer dunklen elliptischen Form suchen.



Abbildung 2: Eyetracking Set-up in der Leitstelle

4 Begründung und Konkretisierung von Forschungsfragen, Forschungsdesign, Zuordnung von Forschungsfragen zu Publikationen und Zusammenfassung der Publikationen

4.1 Begründung der Forschungsfragen

4.1.1 RQ 1 – Demografie, Technikbereitschaft und Team-Arbeit-Kontext

Zu **Begründung der ersten Forschungsfrage (RQ 1)** ist zu erklären, dass zu Beginn dieser Arbeit kaum systematisierte Daten zur Zielgruppe bestehen, besonders nicht solche, die Arbeitskontexte der Leitstellendisponent:innen adressieren. Dies führt zu dem Problem, dass Forschungs- und Entwicklungsziele nur unscharf operationalisiert werden können, beispielsweise durch SMART-Kriterien (Specific, Measurable, Achievable, Realistic, and Timed). Diese Unschärfe liegt nicht nur darin begründet, dass kein Ist-Stand bekannt ist, mit dem etwaige Fortschritte in Bezug auf die Disponent:innen verglichen werden könnten, sondern auch darin, dass die Bedürfnisse der Leitstellendisponent:innen hinsichtlich einer Entscheidungsunterstützung unbekannt sind, was einer menschenzentrierten Entwicklung entgegensteht. Dies ist besonders frappierend, da in den „einschlägigen Wissenschaften“, wie Raehlmann (2017, S. 126) schreibt, Konsens darüber herrscht, dass Technik ein soziales Konstrukt ist. Aus Engineering-Perspektive stellen Göransson et al. (2003, S. 116) „user focus“ und „active user involvement“ auf die beiden Spitzenplätze der anwenderzentrierten Systementwicklung. Folgt man der Prämisse der sozialen Konstruktion von Technik und der Notwendigkeit einer Anwender:innenzentrierung, sind folgerichtig auch die sozialen Kontexte und Bedürfnisse der Anwender:innen zu analysieren, um den daraus resultierenden Bedarf menschenzentriert für den Entwicklungsprozess beschreiben zu können.

4.1.2 RQ 2 – Überzeugungssystem der Leitstellendisponent:innen

Zur **Begründung der Forschungsfrage 2 (RQ 2)** ist zu bemerken, dass durch die Klärung von RQ 1 wichtige Erkenntnisse zu den Disponent:innen zu erwarten sind, jedoch durch die quantitative Forschung keine Erklärungen zu Überzeugungen, Haltungen und Einstellungen von in Deutschland tätigen Leitstellendisponent:innen gegenüber KI-Systemen in Leitstellen möglich sein werden, worüber zum Zeitpunkt dieser Arbeit keine belastbaren wissenschaftlichen Daten existieren. Diese werden jedoch benötigt, wie auch Blomberg et al. (2021) aufzeigen. In ihrer randomisierten klinischen Studie untersuchten sie in der Leitstelle Kopenhagen den Effekt eines Machine-Learning-Verfahrens zur schnellen Erkennung von Herz-Kreislauf-Stillständen während des Notrufdialogs. Dabei berichten sie über robuste Empfehlungen des Systems mit hoher Sensitivität, jedoch schlechter Compliance der Disponent:innen. Blomberg et al. (2021, S. 7) stellen fest, dass die Disponent:innen hätten 54 Herz-Kreislauf-Stillstände erkennen können, wenn sie der Empfehlung des Systems gefolgt wären. Dieses folgenreiche Ablehnen von richtigen Vorschlägen des Systems zeigt die hohe Brisanz der Forschungsfrage und ebenso die enorme Bedeutung für Menschenleben. Dies ist darin begründet, so folgern auch Blomberg et al. (2021, S. 8), dass zusätzlich Überlegungen zu Entscheidungsunterstützungssystemen (Decision Support Systems, DSS) in Leitstellen vor der Implementierung angestrengt werden müssen, besonders solche, die das Vertrauen der Disponent:innen in das System bestärken. Ein DSS ist ein computergestütztes Informationssystem, das die Disponent:innen bei komplexen Problemstellungen und Entscheidungsprozessen, z. B. der Alarmierung des „richtigen“ Rettungsmittels außerhalb von Routinefällen, durch die Bereitstellung relevanter Daten, Analysen und Modelle unterstützt, beispielsweise durch die Miteinbeziehung der Echtzeitverkehrslage. Da jedoch weder Erkenntnisse zu den Vorbehalten, Ängsten und Befürchtungen allgemein vorhanden sind noch zu spezifischen Bedenken, wie z. B. „Alarmierungsmüdigkeit“ bei vielen falsch positiven Hinweisen des Systems, besteht hier eine nennenswerte und hochrelevante Lücke, die mit der Exploration der Überzeugungssysteme geschlossen werden kann.

4.1.3 RQ 3 – Aufmerksamkeitsverteilung der Disponent:innen beim Notrufdialog

Zur **Begründung der Forschungsfrage 3 (RQ 3)** ist darauf aufmerksam zu machen, dass RQ 2 wichtige Erkenntnisse zu den Überzeugungssystemen generieren wird, jedoch keine Einblicke in die tatsächlichen arbeitspraktischen Prozesse, wie vor allem die Interaktion der Disponent:innen mit „der Maschine“, also dem ELS und der Hardware, am Leitstellenarbeitsplatz zu erwarten sind, die über subjektive Beschreibungen hinausgehen.

Diese Einblicke sind jedoch als Vorannahmen für die Entwicklung von DSS essenziell. DSS können die Disponent:innen bei der Entscheidungsfindung unterstützen, indem sie relevante Parameter (z. B. Verkehrssituation, Wetter etc.) in Echtzeit überwachen und eine robuste Vorhersage dieser Parameter ermöglichen (Reuter-Oppermann & Richards, 2019, S. 270). Reuter-Oppermann und Richards (2019, S. 270) fassen das Ziel der DSS wie folgt zusammen: „The overall aim is to help people make informed decisions with the right information and insights at the right time and at the right place, i.e. to the right person.“ Reuter-Oppermann und Richards heben ibidem jedoch auch hervor, dass weiterer Forschungsbedarf hinsichtlich der Gestaltung der grafischen Bedienoberfläche eines DSS und der Warnmeldungen besteht. Ähnliches berichten Franklin et al. (2017), auf die auch Reuter-Oppermann und Richards (2019) referenzieren, für den Bereich der Notaufnahmen, indem sie den hohen Stellenwert verfügbarer Echtzeitdaten für die zeitkritische klinische Entscheidungsfindung hervorheben. Yigitbasioglu und Velcu (2012, S. 56) elaborieren die Perspektive des hohen Stellenwertes mit Hinweisen zum Design von dashboard-basierten DSS. Dabei beschreiben sie universelle Features, die in jedem Fall vorhanden sein sollten. So sollte beispielsweise das Dashboard auf einem einzigen Bildschirm visualisierbar sein und intuitiv erkennbare Symbole enthalten, die zwar ausreichende Salienz gewährleisten, jedoch ohne zu einer optischen Reizüberflutung zu führen. Yigitbasioglu und Velcu (2012, S. 56) heben jedoch auch hervor, dass das Design der DSS zwingend an den Kontext der Nutzung gebunden sein muss.

Es scheint also die Hypothese begründbar, dass DSS zwar auch in den Leitstellen hilfreich sein können und bereits generische Hinweise für das Design eines DSS vorliegen, jedoch noch keine kontextspezifischen Designhinweise für Leitstellen vorliegen, obwohl diese Notwendigkeit besteht. Weiterhin ist auf die hohen mental-informativischen Belastungen der Leitstellendisponent:innen hinzuweisen. Dies unterstreichen auch Herbig und Müller (2016) mit empirischen Daten einer Gefährdungsbeurteilung einer deutschen ILS. Die Autor:innen kommen zu dem Schluss, dass die Disponent:innen vor enormen kognitiven Anforderungen stehen, da sie schnell und präzise Informationen verarbeiten und Entscheidungen treffen müssen. Sie arbeiten in emotional aufgeladenen Situationen, wie bei der telefonischen Anleitung von Reanimationen, und benötigen umfangreiche Kenntnisse in verschiedenen Fachbereichen wie Medizin und Feuerwehrtechnik. Der Arbeitsdruck ist hoch, da die Arbeitsquantität stark variiert und Spitzenzeiten bis zu neun Aufgaben pro Minute umfassen können. Die Ergebnisse der Studie von Herbig und Müller (2016) zeigen eine hohe Arbeitsbelastung der Disponent:innen, die durchschnittlich 0,46 Aufgaben pro Minute

bearbeiten. Während Spitzenzeiten müssen sie bis zu neun Aufgaben pro Minute bewältigen. Die hohe Arbeitsbelastung geht mit signifikant erhöhten Blutdruckwerten und einer geringeren Herzratenvariabilität einher, was auf chronischen Stress hinweist und das kardiovaskuläre Risiko erhöht. Auch die Arbeitsunfähigkeitsrate der Disponent:innen war signifikant höher. Die Studie betont die Notwendigkeit arbeitsorganisatorischer Veränderungen, um die Belastungen der Disponent:innen zu reduzieren. Oldenburg et al. (2014) konnten ähnliche Ergebnisse berichten. Weitere Arbeiten beschreiben ein höheres Risiko für posttraumatische Belastungserkrankungen von Rettungsfachpersonal (Berger et al., 2012).

Diese Erkenntnisse legen nahe, den Status quo der Leitstellenarbeit tiefergehend zu beforschen und Areale mit hoher visueller Aufmerksamkeit und hoher Wichtigkeit versus Areale mit redundanten und distrahierenden Informationen zu identifizieren, um a.) eine Reduktion der mental-informativischen Anforderungen durch technische und organisatorische Maßnahmen herbeiführen zu können, aber auch um b.) daraus generelle Aussagen zum Design von spezifischen DSS für die Leitstellen treffen zu können.

4.2 Konkretisierung, Begründung und wissenschaftstheoretische Verortung des Forschungsdesigns

4.2.1 RQ 1 – Demografie, Technikbereitschaft und Team-Arbeit-Kontext

Das **Forschungsdesign zu Forschungsfrage 1** lässt sich nicht nur durch die hohe praktische Relevanz quantitativer Verfahren der Sozialforschung, in der wiederum standardisierte Fragebögen als typisches Messinstrument zu sehen sind, rechtfertigen (Steiner & Benesch, 2018, S. 47) noch durch die generell dominierende Stellung und populären Einsatz des Fragebogens (Döring et al., 2015, S. 398) innerhalb der Sozialwissenschaften. Vielmehr ist hier hervorzuheben, dass im Rahmen der Beantwortung der RQ 1 gleich drei Vorteile, wie sie auch von Döring und Bortz (Döring et al., 2015, S. 398) beschrieben werden, der Fragebögen als selbstadministrierte schriftliche Befragungsmethode ausgespielt werden können: Erstens können Aspekte des subjektiven Erlebens und privaten Verhaltens erfasst werden, die nicht direkt beobachtbar sind. Zweitens zählt der selbstadministrative Charakter auf das Nebengütekriterium quantitativer Forschung der Forschungsökonomie ein, besonders bei den hier vorliegenden Online-Fragebögen, die bereits für die weitere Analyse optimiert sind, z. B. durch entsprechende Kodierung der Antwortoptionen. Die Forschungsökonomie gewinnt nicht nur, aber besonders in öffentlich-geförderten Projekten, durch den Einsatz „öffentlicher“ und somit nicht zuletzt steuerfinanzierter Mittel, an Bedeutung. Drittens ist die

Fragebogensituation an sich diskreter und anonymer als eine Interviewsituation, die zur Transkription vielleicht zusätzlich noch aufgezeichnet wird, womit auch ein entsprechend sensibel zu handhabendes Datenartefakt entsteht. Dies trifft auf einen völlig anonymen Fragebogen in erhöhtem Maße zu. Nachteilig ist sicher, dass umfangreiche und komplexe Antworten nicht durch einen Fragebogen zu erhalten sind. Dieser Umstand erscheint jedoch verschmerzbar, da es zunächst darum geht, neben soziodemografischen Informationen Einblicke in Technikbereitschaft und bereits kodifizierte Arbeitskontexte zu erhalten (TAKAI). Außerdem wird der eng gesetzte Scope durch RQ 2 ausgeweitet, wie im nächsten Abschnitt ausgeführt wird.

4.2.2 RQ 2 – Überzeugungssystem der Leitstellendisponent:innen

Zur Begründung des **Forschungsdesigns von Forschungsfrage 2** sind v. a. die Ziehung der Stichproben, die Strukturierung des Leitfadens und die Analysestrategie zu erörtern.

Zur Ziehung der Stichproben wird sowohl auf Fokusgruppen als auch auf Expert:innen- bzw. Spezialist:innen-Interviews zurückgegriffen. Die Wahl der Interviewform richtet sich nach dem aktuellen Stand der Theorieentwicklung. So wird diese Entwicklung mit einem Fokusgruppen-Interview initiiert, um bereits aus der regen Diskussion mehrere Stoßrichtungen extrahieren zu können.

Vorwegnehmend ist zu erwähnen, dass die „Übersetzung eines Erkenntnisinteresses in Interviewfragen nur unzureichend durch methodische Regeln unterstützt ist und weitgehend dem Geschick der Interviewer:innen überlassen ist (Gläser & Laudel, 2010, S. 115). Die hier gewählte semistrukturierte Methode des Fragebogens soll die bereits vorhandenen quantitativen Daten ergänzen und ermöglichen, auf unerwartete Themen und Nuancen einzugehen, die in Fragebögen nicht erfasst werden können. Die Leitfragen wurden von Interview zu Interview an den jeweiligen Stand der Theorieentwicklung angepasst und berücksichtigen die Kriterien nach Hopf, nämlich Reichweite, Spezifität, Tiefe und personalen Kontext (Hopf, 1978, S. 99-101 nach Gläser & Laudel, 2010, S. 114).

Die Analyse der Interviews mittels Grounded Theory bietet sich bei dieser theoriegenerierenden Frage besonders an, da es bezüglich der RQ 2 noch keine belastbare Literaturbasis gibt und menschliche Haltungen exploriert werden müssen. Außerdem liegen weitere Vorzüge der GTM in der besonderen Eignung für die Entwicklung von Theorien über Verhaltensmuster und Einstellungen, die Untersuchung von Ängsten und deren Bewältigungsstrategien sowie die Erforschung sozialer Prozesse und Verhaltensweisen in den Sozialwissenschaften. Dies wurde bereits in Kapitel 3.3.2 expliziert.

4.2.3 RQ 3 – Aufmerksamkeitsverteilung der Disponent:innen beim Notrufdialog

Die **Begründung des Forschungsdesigns von Forschungsfrage 3** basiert auf der Prämisse, dass die Möglichkeiten, die visuelle Aufmerksamkeit, die hier im Zentrum steht, nachzuverfolgen (zu tracken), limitiert sind und im Wesentlichen auf Mousetracking und Eyetracking begrenzt sind. Während die Performanz von ET versus MT bereits vielfältig beforscht wurde (Aviz et al., 2019; Egner et al., 2018; Oyekoya & Stentiford, 2005; Rivollier et al., 2021) und generelle Aussagen aufgrund der unterschiedlichen Settings und Endpunkte nur mit Einschränkungen zu führen sind, ist die Performanzperspektive in der hier vorliegenden Arbeit nicht berücksichtigt. Dies ist darin begründet, dass MT entsprechender Software bedarf, um die Bewegungen des Mauszeigers zu registrieren, auf den IT-Instanzen der Leitstellen, die den KRITIS zuzurechnen sind, darf externe Software nur nach erheblichem Aufwand nach notwendigen Sicherheitsprüfungen installiert werden. An diesem Prozess sind neben den IT- und Datenschutz-Spezialist:innen mindestens die Führungskräfte der Leitstelle, der Träger und die ministerielle Aufsichtsbehörde beteiligt, was mit dem Nebengütekriterium der Forschungsökonomie nicht zu vereinbaren ist, wenn Alternativen zur Verfügung stehen. Mit der Eyetracking-Technologie steht jedoch eine solche Alternative zur Verfügung, die sich in den angewandten Wissenschaften auch zunehmender Beliebtheit erfreut (Gegenfurtner et al., 2018) und hinsichtlich Vor- und Nachteilen sowie Gütekriterien und Robustheit der Messung gut beforscht ist, beispielsweise durch Hessels und Hooge (2019), van Renswoude et al. (2018) und Sharafi et al. (2015).

In der konkreten Umsetzung werden vor allem Third- und Fourth-Order-Data (Jacob & Karn, 2003; Karn et al., 1999) synthetisiert, um über sog. Blickpfade die visuelle Aufmerksamkeit nachvollziehen zu können. Über Attentionmaps und Heatmaps können diese Blickpfade analysiert werden. So entsteht ein exaktes Bild, das Aufschluss über die Informationsverarbeitung inklusive Aufgabenpriorisierung und Reihenfolge gibt. Auf diese Weise wird ermöglicht, dass in der zukünftigen Entwicklung von Leitstellensoftware grafische Elemente hervorgehoben werden oder in den Hintergrund treten sowie, dass Bildschirme bedürfnisgerecht angeordnet werden, um Informationen intelligent und kontextsensitiv zu visualisieren und um kognitive Ressourcen zu schonen.

Eine ergänzende Elaboration der wissenschaftstheoretischen Verortung ist mit den Ausführungen zur Eye-Mind-Hypothese und zur Cognitive-Load-Theorie in Kapitel 5.1.3 zu finden.

Bei der Umsetzung ist das Setting der Leitstellen zu berücksichtigen. Meist sind mehrere Bildschirme nebeneinandergeschaltet, teilweise auch zweireihig übereinander. In einigen Leitstellen wird dies noch um weitere, größere Bildschirme an einer großen Wand ergänzt. Somit beschränkt sich das Eyetracking aus technischer Sicht auf mobile Systeme („ET-Brillen“).

4.3 Zuordnung der Publikationen zu den Forschungsfragen

Tabelle 3: Zuordnung der Publikationen zu den Forschungsfragen setzt die im Rahmen der Dissertation eingereichten Publikationen (P1 bis P6) in einen Zusammenhang mit den Forschungsfragen.

Tabelle 3: Zuordnung der Publikationen zu den Forschungsfragen

Nr.	Titel	Zuordnung zu den Forschungsfragen
P1	Technology commitment of emergency medical service practitioners and dispatchers	Forschungsfrage 1: In diesem Artikel wird einerseits die Zusammensetzung der Zielgruppe, andererseits aber auch die Technikbereitschaft von Leitstellendisponent:innen und anderen Anwender:innen adressiert.
P2	Komplexität und Adaptionserfordernisse an der Nahtstelle Leitstelle-Rettungsdienst-Feuerwehr: Eine validierte Fragebogen-analyse	Forschungsfrage 1: Beide Publikationen greifen als „spezielles Merkmal“, wie auch in Forschungsfrage 1 formuliert wurde, den Team-Arbeit-Kontext, der Leitstellendisponent:innen und weiterer Teams auf, die an der rettungsdienstlichen Versorgung teilnehmen.
P3	Zwischen Hierarchie und geteilten mentalen Modellen: Analyse mentaler Modelle an der Nahtstelle Leitstelle-Rettungsdienst-Feuerwehr	
P4	Mental Models of AI in Emergency Medical Services	Forschungsfrage 2: Diese Publikation fokussiert die individuellen Überzeugungen der Leitstellendisponent:innen und weiterer Anwender:innen hinsichtlich KI-Anwendungen in Leitstellen.
P5	Towards Human-AI Interaction in Medical Emergency Call Handling	Forschungsfrage 3: Bei dieser Arbeit werden Grundannahmen der Mensch-KI-Interaktion bei der Notrufannahme und eine Evaluationsdesign für die Feldforschung in der Leitstelle beschrieben. Die Erkenntnisse dienen als Grundlage für die nächste Publikation (P6).
P6	Eye-tracking the Lifeline: Emergency Dispatchers' Visual Attention in Time-Sensitive Calls	Forschungsfrage 3: Dieser Artikel beschreibt die visuelle Aufmerksamkeit von Leitstellendisponent:innen beim Notrufdialog.
Anm.: Ergänzend gibt Anlage 8 einen detaillierteren tabellarischen Überblick über alle Publikationen, inklusive den jeweiligen Forschungsfragen, Stichproben, Methoden, Ergebnissen und Fazits.		

4.4 Zusammenfassung der Publikationen

4.4.1 P1 - Technology commitment of emergency medical service practitioners and dispatchers

Ziel der Publikation ist die Beschreibung der Technikbereitschaft der Leitstellendisponent:innen und weiteren Anwender:innen, um durch die spezifischen Antworten passgenauer auf die Zielgruppe eingehen zu können, beispielsweise mit anschließenden Trainings- und Unterstützungsangeboten.

Hinsichtlich der **Methodik** wurde aufgrund der allgemeinen Ausrichtung des bereits bestehenden Kurzfragebogens zur Technikbereitschaft von Neyer et al. (2016) eine eigene Fragenbatterie entwickelt, die sich auf KI-Anwendungen bezieht, um die Zielgruppe der Leitstellendisponent:innen und Rettungsfachkräfte dahingehend zu befragen. Die weiterentwickelte Fragenbatterie wurde durch eine Fokusgruppe mit Expert:innen aus IT-Berufen und dem Rettungsdienst getestet und verbessert. Der Fragebogen ist online verfügbar (Elsenbast, 2022). Die Antwortoptionen sind in beiden Fällen likertskaliert von 1 bis 5 und entsprechen den Labels „stimmt gar nicht“, „stimmt wenig“, „stimmt teilweise“, „stimmt ziemlich“ und „stimmt völlig“. Der Fragebogen wurde zwischen Oktober 2021 und März 2022 online mit dem Umfragetool LimeSurvey [Version 5.5.0] an die Zielgruppe adressiert. Um einen Vergleich zwischen den Akteur:innen der Notfallversorgung zu ermöglichen, wurden außer den Disponent:innen weitere Berufsgruppen (z. B. Notärzt:innen und Rettungsfachpersonal) befragt.

Bezug nehmend auf die **Ergebnisse** lässt sich festhalten, dass der Fragebogen auf Grundlage der berechneten Gütekriterien (Cronbachs Alpha, McDonalds Omega, konfirmatorische Faktorenanalyse) als valide bezeichnet werden kann. Insgesamt waren 510 Antworten zu verzeichnen. Das Durchschnittsalter der Stichprobe (N = 510) liegt bei 38,9 Jahren (SD = 12,6); n = 354 (69,4 %) sind männlich, 125 (24,5 %) weiblich und n = 4 (0,8 %) sind nicht-binär; n = 68 (13,3 %) sind Notärzt:innen, n = 199 (39,0 %) Leitstellendisponent:innen und n = 184 (36,1 %) Notfallsanitäter:innen. Abbildung 3: Boxplots - Technikbereitschaft zeigt die Ergebnisse nach Gruppen.

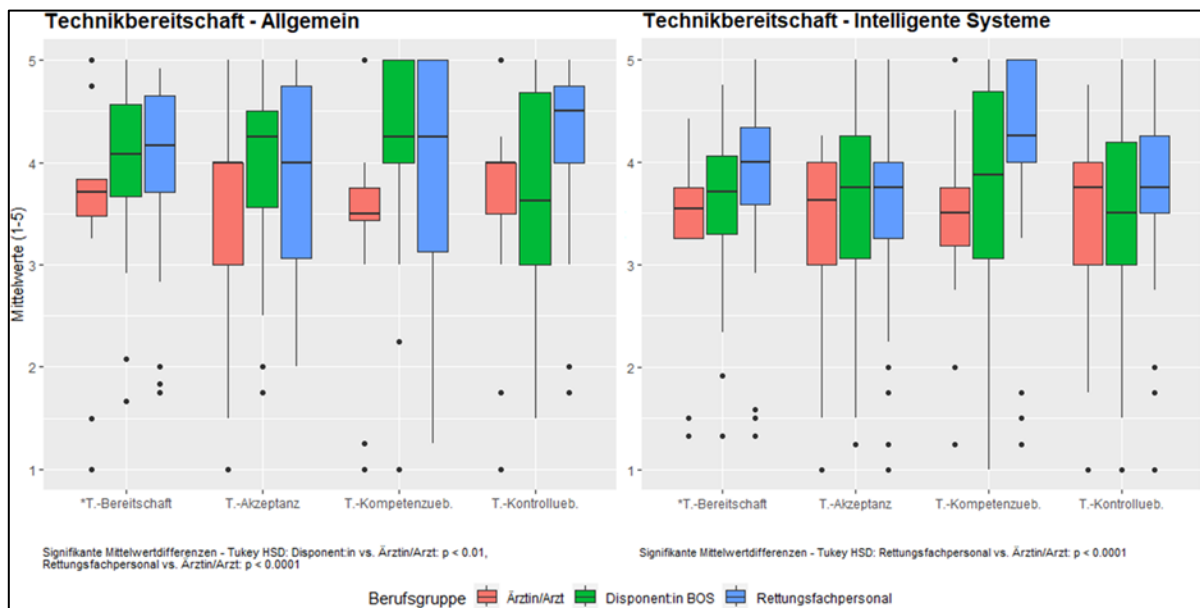


Abbildung 3: Boxplots - Technikbereitschaft

Fazit: Auf Grundlage dieser Daten sind grundsätzlich gute Vorbedingungen für technologische Weiterentwicklungen in Leitstellen und im Rettungsdienst zu erwarten. Ausreißer, also Menschen mit besonders niedriger Technikbereitschaft, dürfen jedoch bei der Entwicklung und Integration neuer Technologien nicht auf der Strecke bleiben. Die Ergebnisse dieser Studie sind spezifisch auf Technikbereitschaft ausgerichtet und erlauben darüber hinaus keine breitere Einschätzung von Arbeitskontexten, wie z. B. der Wahrnehmung von Komplexität oder geteilten mentalen Modellen. Diese Lücke wird mit dem Team-Arbeit-Kontext-Inventar geschlossen (siehe P2 und P3). In Folgeforschungen wäre es wünschenswert, die Repräsentativität durch Erhöhung der Stichprobengröße in Bezug auf die einzelnen Berufsgruppen und ggf. probabilistische Stichproben zu erhöhen. Außerdem sind die expliziten Gründe für niedrige Technikbereitschaft und aversive Haltungen gegenüber KI-Anwendungen durch qualitative Studien zu explorieren. Letzteres adressiert P3.

4.4.2 P2 und P3 - Team-Arbeit-Kontext

Ziel der Publikationen 2 und 3 ist, analog zur Publikation 1, die Beschreibung der Arbeitskontexte der Leitstellendisponent:innen und weiteren Anwender:innen, um durch die spezifischen Antworten passgenauer auf die Zielgruppe eingehen zu können und zielgerichtete Hinweise auf den Funktionsumfang von DSS zu erlangen, die bei der Entwicklung von DSS zu priorisieren sind.

Hinsichtlich der **Methodik** fokussiert das Paper den Einsatz des Team-Arbeit-Kontext-Analyse-Inventars (TAKAI). Die TAKAI-Items waren mit einer sieben-stufigen Likert-Skala zu beantworten. Beispielitems und Konzeptualisierung sind im Forschungsdatenrepositorium und in der Studie von Hagemann et al. frei einsehbar (Elsenbast, 2023; Hagemann et al., 2011). Durch bereits etablierte Trennwerte für die Mittelwerte können kritische Einzelitems und Konstrukte identifiziert werden. Außerdem wurden soziodemografische Parameter erhoben. Der Fragebogen enthielt insgesamt 73 Fragen, die sich auf die Dimensionen Komplexität, Kontextkriterien, Adaptionserfordernisse und geteiltes mentales Modell beziehen. Eine detaillierte Durchsicht des Fragebogens ist über die Online-Materialien (Elsenbast, 2023) oder über die Studie von Hagemann et al. (2011) möglich. Um einen Vergleich zwischen den Akteur:innen der Notfallversorgung zu ermöglichen, wurden weitere Berufsgruppen adressiert (siehe 6.4.1).

Bei den Ergebnissen lässt sich feststellen, dass der Fragebogen aufgrund der berechneten Gütekriterien (Intraklassenkorrelation, Multitrait-Multimethod-Methode) als valide bezeichnet werden kann. Das Durchschnittsalter der Stichprobe (N = 490) beträgt 40,6 Jahre (SD = 11,4); n = 361 (74,4 %) sind männlich, 116 (23,9 %) weiblich und n = 8 (1,7 %) sind nicht-binär; n = 48 (9,8 %) sind Notärzt:innen, n = 39 (8,0 %) Berufsfeuerwehr-Angehörige, n = 43 (8,8 %) Pflegefachkräfte, n = 185 (37 %) Leitstellendisponent:innen und n = 121 (24,7 %) Notfallsanitäter:innen. Die Berufserfahrung im angegebenen Beruf liegt im Durchschnitt bei 16,4 Jahren (SD = 10,9). Detaillierte Ergebnisse sind dem Anhang zu entnehmen (Anlage 4).

Fazit: Strategien und Technologien, die komplexitätsreduzierend wirken, beispielsweise Entscheidungsunterstützungssysteme und Automatisierungen, aber auch interoperable Informationssysteme, sind aufgrund der hohen Komplexitäts- und Intransparenz-Wahrnehmung weiterzuentwickeln und zu integrieren. Trainings sind aufgrund der interprofessionellen Unterschiede speziell auf die jeweiligen Adressat:innen spezifisch auszurichten. Der reine Verweis auf Trainings aus dem Luftfahrtbereich oder den Begriff „gute Kommunikation“ ist hier nicht ausreichend. Auswirkungen prozessualer Änderungen, neue (KI-)Systeme oder die Einbindung telenotärztlicher Konsultationen in die Leitstellenprozesse, sind auch hinsichtlich der Auswirkung auf die HRTs und deren Arbeitskontexte zu beurteilen. In der Folgeforschung sind eine Ausweitung der Stichprobe und eine Veränderung der Rekrutierungsstrategie zu erwägen, analog zu P1.

4.4.3 P4 - Mental Models of AI in Emergency Medical Services

Das **Ziel** dieser Studie ist es, die Überzeugungssysteme (siehe Kapitel 3.2.3) der Disponent:innen in deutschen BOS-Leitstellen hinsichtlich der Nutzung von KI-basierten Entscheidungsunterstützungssystemen (DSS) zu untersuchen, um den Bedarf für gezielte Maßnahmen zur Akzeptanzsteigerung abschätzen und bereits erste Vorschläge dazu machen zu können. Somit wird mit dieser Studie auch der Schlußschluss zu den Erkenntnissen aus P2, aber besonders zu denen aus P1 geschaffen.

Die **Methodik** der Studie basiert auf der GTM (siehe Kapitel 3.3.2). Zunächst wurden Fokusgruppeninterviews durchgeführt, um ein initiales theoretisches Konstrukt zu entwickeln und um Erkenntnislücken zu identifizieren, die in den anschließenden Interviews behandelt werden sollten. Aus logistischen Gründen wurden die Interviewtermine hauptsächlich über Microsoft Teams oder telefonisch durchgeführt. Die Rohdaten wurden auf einer verschlüsselten Serverinstanz gespeichert und in MAXQDA importiert. Insgesamt wurden 8 Interviews mit 31 Personen und einer Gesamtdauer von 619 Minuten zwischen September 2021 und September 2023 durchgeführt und iterativ kodiert. 24 (77,4 %) der Interviewten waren männlich, 7 (22,6 %) weiblich und 0 nicht-binär; 8 (25,8 %) Teilnehmer:innen waren in der Altersgruppe 8-17 Jahre, 17 (54,8 %) in der Altersgruppe 35-52 Jahre und 6 (19,4 %) in der Altersgruppe 53-70 Jahre. Den Teilnehmer:innen wurde Anonymität zugesichert. Weitere methodische Rahmungen finden sich in den Kapiteln 3.2.3 und 3.3.2.

Hinsichtlich der **Ergebnisse** kann festgehalten werden, dass es zu mentalen Modellen der Disponent:innen bisher keine systematisierten und belastbaren Daten gibt, die auf die in Deutschland tätigen Disponent:innen übertragen werden können. Weiterhin ist festzustellen, dass die Arbeitsbelastung der Disponent:innen als hoch eingeschätzt werden muss. Eine weitere Annahme ist, dass KI für die Disponent:innen verständlich sein muss. Des Weiteren hängt die Einstellung gegenüber KI von drei Prädiktoren ab: dem beruflichen Status, dem KI-Wissen und den formalen Qualifikationen. Wenn diese Prädiktoren gering ausgeprägt sind, sind eine hohe Skepsis und eine eher negative Einstellung zu beobachten, was in einer „Blackbox“ endet. Der Begriff „Blackbox“ wird metaphorisch verwendet, um ein System zu beschreiben, das nur in Bezug auf seine Eingaben und Ausgaben verstanden wird, ohne Kenntnis seiner inneren Abläufe. In dieser Studie bedeutet „Blackbox“, dass bestenfalls eine diffuse, nebulöse, volatile und manchmal sogar dystopische Sicht auf KI-Methoden besteht.

Konkrete Implementierungsideen sind selten. Wenn trotz weniger ausgeprägter Prädiktorvariablen eine Reflexion über den Arbeits- und Systemkontext (z. B. rechtliche, politische, gesundheitsökonomische Einflüsse, ethische Überlegungen usw.) erfolgt, ändert sich das Bild zu einer positiven Einstellung. Positive Voreinstellungen führen zu einer differenzierteren Reflexion und zu einer praxisorientierten Beschreibung von Chancen, Möglichkeiten, Risiken, Gefahren und Bedrohungen. Ein hohes Maß an sozialwissenschaftlicher Expertise scheint wichtig für eine tiefgehende Analyse von Ängsten und Vorurteilen, aber das bedeutet nicht, dass nur Absolvent:innen der Sozialwissenschaften eine solche Analyse durchführen sollten.

Fazit: Allgemein wird angenommen, dass eine hohe mentale Belastung in Leitstellen herrscht und dass KI die Disponent:innen, die unter Zeitdruck, Informationsdefiziten und Unsicherheit komplexe Aufgaben bewältigen müssen, unterstützen kann. Dazu müssen KI-Systeme verständlich sein. Außerdem konnten Unterschiede in den Überzeugungssystemen von Disponent:innen zu anderen Expert:innen sowie Prädiktoren für eine negative Haltung gegenüber KI in Leitstellen identifiziert werden. Folgende Maßnahmen können abgeleitet werden: Kurzfristig bedarf es einer signifikant intensivierten Diskussion und Informationsvermittlung über KI innerhalb der Leitstellen. Diese Initiativen müssen sich direkt an die Disponent:innen richten und dürfen sich nicht ausschließlich auf hochrangige wissenschaftliche Fachzeitschriften beziehen. Mittelfristig müssen kompetenzbasierte Lernziele zu KI-Methoden in die Ausbildungsprogramme für Disponent:innen und Rettungsfachpersonal integriert werden. Angesichts der unterschiedlichen Zugangsvoraussetzungen, Ausbildungsinhalte und Ausbildungsdauern für die Tätigkeit als Disponent:in in den verschiedenen Bundesländern (Trautmann et al., 2022, S. 80–94), sollte dies mit einheitlichen Standards, wie einem Erwartungshorizont für die Abschlussqualifikation, einhergehen. Außerdem müssen KI-Systeme erklärbar und möglichst intuitiv bedienbar sein, was einen Schwerpunkt in der Evaluation darstellen sollte. P5 greift speziell diese Erkenntnis auf. Mittel- und langfristig sind die Wechselwirkungen zwischen KI und Disponent:innen durch empirische Feldstudien zu erforschen, um die spezifischen Gründe für Ablehnungen in konkreten Fällen zu explizieren. Dieses Unterfangen erfordert einen interdisziplinären Ansatz und die Zusammenarbeit mit Fachkräften in den Leitstellen. Folgeforschung sollte die generierte Theorie an konkreten Praxisbeispielen in den Leitstellen, in denen KI-Verfahren eingesetzt werden, testen.

4.4.4 P5 - Towards Human-AI Interaction in Medical Emergency Call Handling

Ziel der Studie ist die Entwicklung eines menschenzentrierten Designs von sog. hybriden KI-Systemen in der Notrufannahme, um eine möglichst hohe Passung zwischen erklärbarer bzw. verstehbarer Maschine und menschlichen Verarbeitungs- und Arbeitsprozessen zu schaffen. „Design“ kann hier synonym zu „Konzept“ verstanden werden (Kerres, 2018, S. 232). Dies inkludiert auch einen Evaluationsansatz, der sich auf das hybride KI-System bezieht.

Die **Methodik** dieser Arbeit basiert auf zwei Säulen. Erstens wurden die Leitlinien für Mensch-KI-Interaktion (Amershi et al., 2019) von den Autor:innen auf die praktischen Bedarfe der Leitstellen transferiert, um die Softwareartefakte für ein hybrides KI-Entscheidungsunterstützungssystem zur Notrufabfrage und Einsatzmittelalarmierung zu charakterisieren. Zur Begrifflichkeit „hybride KI“: Ein grundlegender Bestandteil einer solchen KI ist die Schnittstelle, die eine Mensch-KI-Interaktion ermöglicht, beispielsweise via Chatbot. Im Laufe dieser Interaktion lernt der Mensch von der KI und die KI kann Datenpunkte verknüpfen, um ihre Leistung im Laufe der Zeit zu verbessern. Softwareartefakte sind als Elemente einer Roadmap zu verstehen, die die partizipative Entwicklung in einem Workflow ermöglichen. Zweitens wurde darauf aufbauend unter Berücksichtigung von etablierten Evaluationsstandards (DeGEval – Gesellschaft für Evaluation, 2016) ein grundlegendes Evaluationsdesign für KI im Leitstellensetting definiert, um eine menschenzentrierte Bewertung der Artefakte vornehmen zu können.

Als **Ergebnis** kann auf den im Rahmen dieser Arbeit definierten iterativen Entwicklungsprozess verwiesen werden, der aus Design, vorläufiger Evaluation, Implementierung und Evaluation besteht und sich in puncto Evaluation in fünf Schritte aufgliedern lässt. Diese sind (1) Aufklärung und Einwilligung der Versuchspersonen, (2) Applikation und Kalibrierung von Sensoren (z. B. Eyetracker), (3) Entspannungssequenz, (4) Erhebungsphase sowie (5) Ausfüllen eines Fragebogens, um beispielsweise die Einordnung von subjektiven Eindrücken der Disponent:innen vornehmen zu können. Die erforderlichen Artefakte für eine Mensch-KI-Interaktion bei der Bearbeitung von Notrufen zielen darauf ab, den Disponent:innen bzw. Führungskräften geeignete Vorschläge für die konkreten Entwicklungsschritte geben zu können: (1) „Kläre die benötigten Funktionen des Systems im Voraus“; (2) „Kläre die benötigte Leistung des Systems im Voraus“; (3) „Lege den richtigen Zeitpunkt für die Interaktion des Systems mit den Disponent:innen fest“; (4) „Stelle kontextuell relevante Informationen

bereit“; (5) „Ermögliche eine effiziente Ablehnung und Korrektur“; (6) „Biete im Zweifelsfall Alternativen an“; (7) „Erkläre die grundlegende Funktionsweise der KI“; (8) „Lerne aus dem Benutzerverhalten und biete Feedbackmechanismen“; (9) „Kläre die Auswirkungen der Benutzeraktionen auf zukünftige Interaktionen“; (10) „Halte Anpassungsmöglichkeiten für die Benutzer:innen überschaubar“. Diese Artefakte wurden als Resultat eines Fokusgruppeninterviews mit rheinland-pfälzischen Leitstellenmitarbeiter:innen entwickelt.

Fazit: Durch den Transfer der Guidelines für Mensch-KI-Interaktion in das Praxisfeld der Leitstellen, konnten konkrete Schlüsselemente für KI-Designs zur Notrufbearbeitung identifiziert werden und somit eine Orientierung für Leitstellen und Entwickler:innen geschaffen werden, um auf hybrider KI basierende DSS für Leitstellen zu etablieren. Nachdem die Grundsätze der Evaluation für DSS im Leitstellenkontext beschrieben worden sind, bestehen Lücken bezüglich der Interaktion von Disponent:innen und grafischer Bedienoberfläche, die für die bedarfsorientierte und menschenzentrierte Entwicklung von DSS wichtig sind. Außerdem ist die subjektive Wahrnehmung der Disponent:innen in Bezug auf die aufgabenbezogenen Anforderungen während der Notrufbearbeitung noch unbeforscht. P6 wird sich dieser beiden Lücken annehmen.

4.4.5 P6 – Visuelle Aufmerksamkeit in Leitstellen der Gefahrenabwehr

Ziel der Studie ist, per Eyetracking Informationen zur visuellen Aufmerksamkeitsverteilung der Disponent:innen und per Fragebogen Daten zur Wahrnehmung der Disponent:innen von aufgabenbezogenen Anforderungen während der Notrufbearbeitung zu erheben, um Hinweise zu erhalten, die es ermöglichen, die Gestaltung von Entscheidungsunterstützungssystemen (DSS) an den tatsächlichen Gegebenheiten der Mensch-Maschine-Interaktion auszurichten.

Die **Methodik** stützt sich auf binokulares Eyetracking, es wurden also beide Augen getrackt. Zwei Leitstellen wurden berücksichtigt. Eine Leitstelle hatte eine Konfiguration mit vier Hauptmonitoren, während die andere zwei Hauptmonitore verwendete. Details sind der Publikation im Anhang zu entnehmen. Zur Erkennung der Blickpunkte als Rohdaten zur Berechnung der Fixationen wurde der mobile Eyetracker „Pupil Labs Core“ (Pupil Labs GmbH, o. D.) verwendet. Es wurden spezifische Interessenbereiche (Areas of Interest, AOIs) in der Analysesoftware [Blickshift Analytics Professional, Version 2024.5.1] ausgewertet. Die AOIs entsprechen jeweils den Monitoren und den Touchscreens (jeweils zwei). Zur Ergänzung der

Eyetracking-Daten wurde ein Online-Fragebogen an die teilnehmenden Disponent:innen adressiert, der soziodemografische Informationen und die Arbeitsbelastung erfasste. Dazu wurde eine verkürzte (ungewichtete) Version des NASA Task Load Index, der sog. Raw-TLX nach Hart und Staveland (1988), mit sechs Dimensionen verwendet: mentale Anforderung, physische Anforderung, zeitliche Anforderung, Leistung, Anstrengung und Frustration. Weiterhin wurden Maßnahmen zur Sicherung der Datenqualität etabliert, beispielsweise eine Kalibrationsmethode für den Eyetracker (Santini et al., 2017) und ein Arbeitsprozess zur Analyse der Eyetracking-Daten (Godfroid & Hui, 2020; Raschke, 2024) inklusive Trennwert für die Accuracy des Trackings (Mean Confidence > 0,7).

Bezüglich der **Ergebnisse** ist zu erwähnen, dass von initial 33 Disponent:innen 30 Datensätze mit einer durchschnittlichen Beobachtungsdauer von 60 Minuten analysiert werden konnten, da in zwei Fällen die Datenqualität unzureichend war und eine Versuchsperson das Experiment auf eigenen Wunsch vorzeitig verlassen hat. Die Eyetracking-Analyse (N = 30) zeigte signifikante Unterschiede in den visuellen Aufmerksamkeitsmustern zwischen den beiden Leitstellenkonfigurationen. In Leitstelle 1 (Vier-Monitor-Setup, n = 15) konzentrierten sich die Disponent:innen intensiv die beiden inneren Monitore, mit minimaler Nutzung des Touchscreens. Leitstelle 2 (Zwei-Monitor-Setup, n = 15) zeigte ein anderes Muster, mit starker visueller Konzentration auf der linken Seite des linken Bildschirms, die sich zur Mitte hin ausbreitete. Beide Setups wiesen Bereiche hoher visueller Fokussierung und Regionen auf, die weitgehend ignoriert wurden. So wurden in Leitstelle 1 der linke Monitor kaum benutzt und der rechte Monitor nur geringfügig mehr. Bemerkenswert ist, dass die Touchscreens in beiden Leitstellen selten genutzt wurden. P 6 quantifiziert detailliert mit Angabe der Fixationsfrequenz und Fixationsdauer der AOIs.

Die Raw-TLX-Ergebnisse zeigten signifikante Unterschiede in der wahrgenommenen Arbeitsbelastung zwischen den beiden Leitstellen. Disponent:innen in Leitstelle 1 berichteten höhere Niveaus mentaler Belastung und Zeitdruck im Vergleich zu denen in Leitstelle 2.

Hinsichtlich der Datenqualität erreichte die Studie eine durchschnittliche Confidence von 0,83 und bei Disponent:innen mit Brille 0,75. Letzteres kann dem häufigeren Auftreten von störenden Reflexionen bei Brillenträger:innen zugerechnet werden. Zuverlässige Erklärungen für fehlende Werte konnten jedoch nicht immer gefunden werden. In unklaren Fällen kann ein Drift, also ein graduelles Verrutschen des Eyetrackers über die relative lange Beobachtungsdauer vermutet werden (Sharafi et al., 2020, S. 7).

Bei dieser Feldstudie konnte ein Verhältnis von Aufnahmedauer zu wissenschaftlicher Bearbeitung von 1:16 dokumentiert werden. Das heißt, dass 60 Minuten Eyetracking-Daten 960 Minuten Arbeitszeit bedürfen, worin Vorbereitung, Durchführung und Nachbereitung des Experiments und die Analyseschritte inkludiert sind. In toto kann also von 480 Stunden Bearbeitungsdauer, exklusive Publikationsleistung, ausgegangen werden. Ein großer Teil der Bearbeitungsdauer ist durch das „Trainieren“ von Objecttracking und Interpolation zu erklären, das immer wieder manuell korrigiert werden musste.

Als **Fazit** lässt sich festhalten, dass eine bloße Erhöhung der Monitoranzahl die Effizienz in Leitstellen nicht zwangsläufig zu verbessern scheint. Stattdessen wird die Notwendigkeit optimierter, auf die Arbeitsabläufe der Disponent:innen zugeschnittener Visualisierungsstrategien deutlich. Es bestehen Hinweise auf die Überlegenheit des Zwei-Monitor-Set-ups in Bezug auf visuelle Aufmerksamkeitsverteilung und selbstberichtete Arbeitsbelastung. Zukünftige Forschung sollte größere, vielfältigere Leitstellen-Setups, standardisierte Aufgabenszenarien und ergänzende Maßnahmen der kognitiven Verarbeitung einbeziehen. Eine longitudinale Perspektive wäre wertvoll, um Veränderungen der Aufmerksamkeitsmuster über die Zeit zu verstehen. Die Ergebnisse unterstreichen die Bedeutung nutzerzentrierter Ansätze bei der Gestaltung von DSS in Leitstellen. Eine kritische Reflexion von Drittvariablen ist jedoch unerlässlich und in Folgeforschungen umzusetzen.

5 Diskussion

5.1 Interpretation

Um eine bessere Orientierung zu gewährleisten, sind die Ausführungen nach den Forschungsfragen 1 bis 3 gruppiert und in Einleitung, Interpretation, Einordnung in den Forschungskontext, kritische Reflexion, Implikationen und Ausblick gegliedert. Ergänzend gibt Anlage 8 einen detaillierteren tabellarischen Überblick über alle Publikationen, inklusive den jeweiligen Forschungsfragen, Stichproben, Methoden, Ergebnissen und Fazits.

Aufgrund der umfassenden ersten Forschungsfrage wurde der Abschnitt zur RQ 1 als einziger in weitere drei Abschnitte zergliedert, um eine einfachere Orientierung zu ermöglichen. Die drei Abschnitte lauten Soziodemografie, Technikbereitschaft und Team-Arbeit-Kontext.

5.1.1 RQ 1 - Soziodemografie, Technikbereitschaft und Team-Arbeit-Kontext

5.1.1.1 Soziodemografie

Einleitung: Die erste Forschungsfrage wird hinsichtlich des Schwerpunkts Technikbereitschaft in P1 (N = 510) und hinsichtlich Team-Arbeit-Kontext in P2 und P3 (jeweils N = 490) adressiert. Bei beiden Untersuchungen handelt es sich um Fragebogenstudien, die sich online primär an Leitstellendisponent:innen, Rettungsfachpersonal, und Notärzt:innen richten.

Interpretation: Zusammenfassend sind hinsichtlich der Soziodemografie der Zielgruppe in beiden fragebogenbasierten Studien (P1, P2/P3) ähnliche Ergebnisse zu verzeichnen, was die Geschlechterverteilung, das Alter und die Beteiligung der Leitstellendisponent:innen angeht. Grundsätzlich ist von eher lebenserfahrenen Disponent:innen auszugehen, die bereits auf mehr- und langjährige Berufserfahrung zurückblicken können. Das Einstiegsalter kann durch die vorher obligatorisch zu erlangende Berufserfahrung in Rettungsdienst und/oder Feuerwehr erklärt werden. Es zeigt sich in puncto Geschlechterverteilung, dass die Leitstellen deutlich männlich besetzt sind. Die Stichprobengröße kann zunächst als zufriedenstellend verstanden werden und in Bezug auf die soziodemografische Zusammensetzung scheint Plausibilität anzunehmen zu sein, wie die Einordnung im nächsten Abschnitt verdeutlicht.

Einordnung in den Forschungskontext: Die Lage und Verteilung von Alter und Berufserfahrung sind vergleichbar mit den Leitstellenstudien von Trautmann et al. (2022, S. 24) und Trautmann und Ballé (2022, S. 13). Die beiden vorgenannten Studien stellen bis dato die teilnehmerstärksten Studien aus dem deutschen Leitstellenbereich dar. Auch im Vergleich mit stichprobenstarken Studien aus dem deutschsprachigen Rettungsdienst, wie beispielsweise der Berufstreuestudie 1 (Hofmann & Macke, 2020, S. 7) und der Berufstreuestudie 2 (Elsenbast et al., 2023, S. 11), können ähnliche Verteilungsmuster festgestellt werden. Lediglich zur Arbeit von Trautmann et al. (2022, S. 24), die mit einem Anteil männlicher Befragungsteilnehmer von 93% einen extrem hohen Männeranteil verzeichnete, bestehen große Unterschiede. Zur extrem abweichenden Geschlechterverteilung in der Studie von Trautmann et al. (2022) lassen sich ad hoc keine Erklärungsmuster explizieren, auch weil dies in einer Studie (Trautmann & Ballé, 2022, S. 24) mit ähnlichem Befragungszeitraum und identischer Zielgruppe nicht mehr gezeigt werden konnte. Die Durchschnittsangaben zum Alter dürfen jedoch nicht über die große Spanne des Altersbereichs hinwegtäuschen. Bezüglich der Personalförderung von altersmäßig heterogenen Belegschaften weisen Kollmann et al. (2020) auf die Herausforderung für Organisationen hin, sowohl jüngere als auch ältere Mitarbeiter:innen zu fördern.

Reflexion: Nachdem amtliche Erhebungen zur Soziodemografie der Zielgruppe fehlen, kann mit Verweis auf die o. g. teilnehmerstarken Studien vorläufig davon ausgegangen werden, dass sich die Stichproben der vorliegenden Studie (P1 bis P3) zu RQ 1 in puncto Soziodemografie mit der Grundgesamtheit decken. Limitierend ist jedoch, dass es sich sowohl bei den Studien im Rahmen dieser Dissertation als auch bei den in diesem Kapitel zitierten Studien um Gelegenheitsstichproben und somit um nicht-probabilistische Stichproben handelt, die durch Selbstselektionsverzerrungen mit der Annahme falscher Populationsverhältnisse verbunden sein können (Döring et al., 2015, S. 306–307). Gelegenheitsstichproben dürfen bei Populationsbeschreibungen nicht überinterpretiert werden, was jedoch ihre Nützlichkeit für Theorieprüfung und Theoriebildung nicht in Abrede stellt (Döring et al., 2015, S. 307).

Implikationen: Die Personalplanung sollte auch lebens- und berufserfahrene Mitarbeiter:innen über 40 Jahren berücksichtigen. Weiterbildungen sind diesem Erfahrungsniveau ebenso anzupassen. Für Jugendliche und junge Erwachsene kann die Bedeutung der sozialen Medien hervorgehoben werden (Bundeszentrale für politische Bildung, 2024), was sowohl die Personalrekrutierung als auch die Informationsbeschaffung durch die Disponent:innen (beispielsweise bei Großschadenslagen) betrifft. Über die Ansprache jüngerer Menschen per Social Media könnte direkt für das technologisierte Feld der Leitstellen sensibilisiert und auch begeistert werden, was ein wichtiger Bestandteil der Nachwuchsgewinnung sein kann.

Die Altersverteilung könnte sich jedoch bei der Einführung eines neuen Ausbildungsberufs „Leitstellendisponent:in“ verändern, was weiter in Folgeforschungen zu berücksichtigen ist. Für populationsbeschreibende Studien zu den Leitstellendisponent:innen, die für Deutschland bislang fehlen, sind probabilistische Stichproben essenziell.

Ausblick: Weitere Forschung zur Gesamtpopulation der Leitstellendisponent:innen ist wünschenswert, da bisher beispielsweise zur Anzahl der in Deutschland tätigen Disponent:innen belastbare Angaben fehlen und die Grundgesamtheit bei ca. 230 Leitstellen mit sehr unterschiedlichen Größen nicht zuverlässig geschätzt werden kann.

5.1.1.2 Technikbereitschaft

Einleitung: P1 (N = 510) ist RQ 1 zuzurechnen und fokussiert den Themenkomplex Technikbereitschaft mit einer Online-Fragebogenstudie und einer Weiterentwicklung der

Kurzskala Technikbereitschaft, die sich gezielt auf KI-Anwendungen in der Leitstelle fokussiert. Die weiterentwickelte Kurzskala wird u. a. per konfirmatorischer Faktorenanalyse (CFA) nach Feststellung der Reliabilitätsparameter validiert.

Interpretation: Die fragebogenbasierte Studie lässt darauf schließen, dass eine hohe allgemeine Technikbereitschaft von Disponent:innen, Rettungsfachpersonal und an der Notfallversorgung beteiligten Ärzt:innen gegenüber Alltagstechnologien vorliegt und eine etwas geringere, aber immer noch gute Technikbereitschaft gegenüber intelligenten Systemen in der Leitstelle. Die in dieser Studie berechneten Gütekriterien (u. a. Cronbachs Alpha und CFA) sprechen für Zuverlässigkeit und Messgenauigkeit des eingesetzten Fragebogens. Die Ergebnisse ähneln denen von Neyer et al. (2012). Dies bedeutet, dass seitens der Endanwender:innen grundsätzlich gute Voraussetzungen für die Entwicklung und Integration intelligenter Systeme im Leitstellenumfeld vorliegen. Jedoch bleibt zu berücksichtigen, dass auch statistische Ausreißer in Richtung geringer Technikbereitschaft zu verzeichnen sind, die wiederum nur ungenügend über die Variablen Alter und Geschlecht erklärt werden können. Hinsichtlich des Zusammenhangs von Technik und Alter heißt dies, dass die Annahme, ältere Disponent:innen seien per se weniger technikbereit als jüngere Disponent:innen, zwar nominell durch die abnehmende Technikbereitschaft zwischen jungen und älteren Altersgruppen gezeigt werden konnte, jedoch die praktische Relevanz fraglich ist. Eine einfache lineare Regression mit Technikbereitschaft als Kriterium und Alter als Prädiktor erwies sich als signifikant ($F(1, 465) = 66,33; p < 0,001$). 12,48 % der Varianz in der Technikbereitschaft können durch die Variation des Alters erklärt werden. Der Regressionskoeffizient der Variable Alter betrug $r = -0,021$ und war signifikant ($p < 0,001$). Demzufolge können wir annehmen, dass für jedes zusätzliche Lebensjahr die Technikbereitschaft im Durchschnitt um 0,021 Einheiten abnimmt. Über 50 Lebensjahre hinweg ist also eine Absenkung der Technikbereitschaft um 1,05 Punkte anzunehmen, was die kritische Diskussion der o. g. praktischen Relevanz begründet. In einer multiplen linearen Regression wurden keine relevanten Effekte für die anderen erfassten Variablen (Berufserfahrung, Geschlecht, Verantwortlichkeit für regionale Leitstelle und Leitstellengröße) gefunden. Insgesamt sind in der Literatur heterogene Ergebnisse zum Zusammenhang zwischen Lebensalter und Technikbereitschaft feststellbar (Czaja et al., 2006; Mollenkopf & Kaspar, 2004; Neyer et al., 2012).

In puncto statistische Ausreißer, also Menschen, die in besonders niedrigem Maße technikbereit sind, ist zu betonen, dass dieser analytische Begriff nicht nur Datenpunkte beschreibt, sondern Menschen und ihre Arbeitsrealität als immanente Dimension der sozialen Teilhabe (Nationale Armutskonferenz, 2017, S. 3).

Einordnung in den Forschungskontext: Strutz et al. stellen nicht auf das Lebensalter oder das Geschlecht ab, sondern auf die tatsächliche Techniknutzung als moderierenden Faktor und schließen, „dass Techniknutzung, also Erfahrung und Fertigkeiten mit digitaler Technik, ein relevantes Kriterium von Technikkompetenz ist.“ (Strutz et al., 2020, S. 7). Es erscheint schlüssig, dass Techniknutzung von vielen weiteren Variablen abhängt, wie beispielsweise der evident zunehmenden Familiarisierung in den verschiedenen Berufsbildern mit digitalen Technologien im Rahmen der fortschreitenden Digitalisierung. Diese Annahme erklärt auch die mitunter signifikanten Unterschiede zwischen den Berufsgruppen in der vorliegenden Studie (P1). In der Gesamtschau bedeutet dies, dass vor allem hinsichtlich des Alters eine Kontextabhängigkeit vorzuliegen scheint und Technikbereitschaft weiterhin von der konkret zu betrachtenden Technologie abhängt. Venkatesh und Davis (2000) zeigen in ihrer Erweiterung des Technikakzeptanzmodells, das auch mit dem Konzept der Technikbereitschaft nach Neyer et al. (2012) als zugrundeliegende Theorie der Ausführungen dieses Abschnitts zu verstehen ist, bereits auf, dass eine Erklärung der Ausprägung von Technikbereitschaft nicht als monokausaler Zusammenhang erfasst werden kann. Vielmehr sind die maßgeblich tatsächliche Nutzung, die Absicht zur Nutzung sowie die Einstellung zur Technologie durch verschiedene Variablen moderiert, in deren Kern die wahrgenommene Nützlichkeit und die wahrgenommene Bedienerfreundlichkeit stehen.

Reflexion: Als Limitation sei auf die bereits erwähnten Verzerrungen durch Selbstselektion hingewiesen. So könnten durch den Online-Modus der Befragung wenig technikbereite Personen nicht teilgenommen haben, was zu einer falsch optimistischen Darstellung der Technikbereitschaft führen würde. Ob dies in der Praxis zu einer Bedrohung der externen Validität führt, wäre jedoch kritisch zu diskutieren, da es wenig wahrscheinlich ist, dass diejenigen, die sehr wenig technikbereit sind, in ausgeprägten technischen Umgebungen wie Leitstellen arbeiten. Weiterhin sind in Folgeforschungen größere Umfänge der einzelnen Berufsgruppen zu berücksichtigen, um die Aussagekraft der Vergleiche zwischen den Gruppen zu erhöhen. Das Konstrukt der Technikbereitschaft erfasst allgemeine Dispositionen, unabhängig davon, ob die Nutzung freiwillig oder verpflichtend ist. Ob ein beruflich bedingter

Verpflichtungscharakter der Techniknutzung zu einem Bruch der Technikbereitschaft führt, kann hier nicht geklärt werden, ebenso wenig wie die expliziten Gründe für niedrige Technikbereitschaft in den Leitstellen.

Implikationen: Die Ausführungen zu den statistischen Ausreißern zeigen die Notwendigkeit auf, Techniknutzung aus einer deontologisch-ethischen Perspektive zu betrachten. Dabei geht es auch darum, wie die in der Studie dargestellten Ausreißer bei einer Implementierung in der Leitstelle an die Technologie herangeführt und etwaige Stereotype abgebaut werden können und ein Gefühl der Nützlichkeit und Relevanz der Implementierung erzeugt werden kann.

Verweisend auf Davis und Venkatesh (2000) ist die wahrgenommene Nützlichkeit ein wichtiger Prädiktor für die Einstellung zur Technologie und somit auch für die Absicht zur Nutzung. Die Nützlichkeit des jeweiligen KI-Verfahrens für die Disponent:innen ist folglich nicht nur ein zentrales Anliegen der Entwicklung und Integration, sondern auch Bestandteil der Kommunikation, die auch zielgerichtet in Richtung der Disponent:innen zu forcieren ist. Dabei ist besonders die berufliche Relevanz herauszustellen. Die zu beantwortende und zu kommunizierende Kernfrage lautet daher: „Welche berufliche Relevanz im Sinne eines Nutzens hat die Technologie für die Disponent:innen?“

Die wahrgenommene Nützlichkeit ist nach Davis und Venkatesh (2000) auch von der Erfahrung, den subjektiven Normen und dem (Produkt-/Marken-)Image abhängig. Dies bedeutet, dass grundsätzlich ein Qualifikationskonzept zu etablieren ist, das einerseits die Erfahrungen der Disponent:innen wertschätzend berücksichtigt, was gleichermaßen auch einen Teil der Notwendigkeit einer partizipativen Entwicklung und Integration anzeigt. Andererseits sind neben den Erfahrungen auch die unterschiedlichen Rollen in den Leitstellen zu berücksichtigen. Das betrifft beispielsweise lokale Gegebenheiten wie die Organisation in Calltaker-Dispatcher-Arrangements, aber auch die verschiedenen Hierarchiestufen – von Disponent:innen über IT-Spezialist:innen und Führungskräfte bis hin zu Betreibern und Trägern. Eventuelle Vorbehalte durch das Image „der KI“ können hier ebenfalls adressiert werden, ebenso wie der exakte Modus der Mensch-Maschine-Interaktion, die Transparenz der KI-Prozesse und weitere Aspekte (z. B. Rechtssicherheit, Datensicherheit etc.). Die Kernfrage lautet hier: „Welche Kompetenzen sind zu welchem Zeitpunkt mit welchen Maßnahmen bei welcher Stakeholdergruppe zu vermitteln?“

Weiterhin ist nicht nur die Ergebnisqualität, sondern auch die Darstellbarkeit der Ergebnisse für die wahrgenommene Nützlichkeit entscheidend (Venkatesh & Davis, 2000). In Bezug auf

die wahrgenommene Nützlichkeit dürfte auch die Geschwindigkeit des DSS ein wichtiger Parameter sein. Außerdem sollten frühe technische Entwicklungen, die „hands-on“ für die Disponent:innen dargestellt werden, darauf ausgerichtet sein, Disponent:innen als Fürsprecher:innen und früher Multiplikator:innen innerhalb der Leitstelle zu gewinnen. Als Kernfrage kann gelten: „Zu welchem Zeitpunkt sind die ersten Ergebnisse an die Zielgruppen der Disponent:innen zu kommunizieren, um eine möglichst hohe Zahl von Fürsprecher:innen zu gewinnen?“

Technikbereitschaft, auch speziell hinsichtlich intelligenter Systeme, scheint sich im Kontext der technikzentrierten Leitstellen als ein Rekrutierungskriterium in der Personalauswahl der Disponent:innen anzubieten.

Ausblick: In Zukunft werden durch die fortschreitende Digitalisierung zielgerichtete Konzepte benötigt, die es ermöglichen, Menschen mit hoher Technikbereitschaft zu bestätigen und Menschen mit niedriger Technikbereitschaft in der Entwicklung nicht zurückzulassen. Die jeweiligen Gründe für eine niedrige Technikbereitschaft und die Auswirkung einer eher unfreiwilligen KI-Anwendung sind vor Ort weiter zu explorieren. Außerdem sind in Folgeforschungen größere Stichproben einzelner Berufsgruppen zu berücksichtigen.

5.1.1.3 Team-Arbeit-Kontext (TAKAI)

Einleitung: P2 und P3 (N = 490) sind ebenfalls RQ 1 zuzurechnen und fokussieren den Themenkomplex Team-Arbeit-Kontext mit einer Online-Fragebogenstudie, die das Team-Arbeit-Kontext-Inventar (Hagemann et al., 2011) einsetzt. Eine Validierung im Kontext dieser Studie erfolgt mit Berechnung der Intraklassenkorrelation und der Interkorrelationen im Heterotrait-Heteromethod-Block.

Interpretation: Bei der Komplexitätswahrnehmung überschreiten alle Berufsgruppen (Leitstellendisponent:innen, Notfallsanitäter:innen, Berufsfeuerwehr-Angehörige, Notärzt:innen, Pflegefachpersonal) den kritischen Wert (≥ 3) der Items, mit Ausnahme der Pflegefachkräfte. Dies deutet auf eine hohe Komplexitätswahrnehmung bei den Angehörigen der HRTs hin.

Bei den Adaptionserfordernissen überschreiten Notfallsanitäter:innen und Leitstellendisponent:innen den Trennwert ($\geq 5,5$) bei der Informationssammlung (MW = 5,56 bzw. 5,59). Dies zeigt die Notwendigkeit der Verbesserung adaptiver Verhaltensmuster an, die ein Team dazu befähigen, sich flexibel an sich verändernde Kontextbedingungen anzupassen

und adäquat darauf zu reagieren (Hagemann, 2021, S. 4). Zu den Adaptionserfordernissen werden beispielsweise Aufgabenverteilung sowie -priorisierung gezählt.

Beim geteilten mentalen Modell heben sich die Pflegefachkräfte von den anderen Gruppen ab, ähnlich die Notärzt:innen beim Shared Team Mental Model. Auffällig ist, dass bei allen Items diese Kategorie die Mittelwerte nahe dem Trennwert (≥ 5) liegen, was als reduzierte Abbildung der Wirklichkeit auf eine heterogene Wahrnehmung der Situation mit dem Risiko von Zielkonflikten hinweist. Die Gruppenunterschiede sind beim Shared Task Mental Model zwischen den Pflegefachkräften und den Berufsfeuerwehr-Angehörigen ($\Delta = 1,23$) am größten, beim Shared Team Mental Model zwischen den Notärzt:innen und den Berufsfeuerwehr-Angehörigen ($\Delta = 0,97$). Die Unterschiede sind bedeutend für die interprofessionelle Zusammenarbeit an den jeweiligen Nahtstellen (z.B. Schockraum und Leitstelle), da hier eine gemeinsame Zielorientierung essenziell ist.

In der Hierarchiewahrnehmung, die den Kontextfaktoren zuzurechnen ist, überschreiten Notfallsanitäter:innen als einzige Gruppe nicht den Trennwert. Die Ursache kann in der Hierarchie selbst liegen. Während Notärzt:innen im Einsatz dem „nicht-ärztlichen Personal“ in medizinischen Fragen Weisungen erteilen können, befinden sich die anderen Berufsgruppen oftmals in der Rolle der Weisungsempfänger. Jedoch herrschen im Hinblick auf den Umfang dieser Weisungsbefugnisse oftmals Unklarheiten und es bestehen Fallstricke, z. B. wenn medizinische und einsatztaktische Aspekte nicht klar abgrenzbar sind (Perthes & Jakob, 2012). Die in dieser Studie berechneten Gütekriterien sprechen für Zuverlässigkeit und Messgenauigkeit des eingesetzten Fragebogens.

Einordnung in den Forschungskontext: Die Ergebnisse sind mit der Studie von Hagemann et al. (2011) vergleichbar. Die Befunde korrespondieren außerdem mit Dörners (2008) Ausführungen zu Komplexität. Dörner identifiziert Merkmale wie Vernetztheit, Intransparenz und Eigendynamik als zentrale Charakteristika komplexer Systeme. Die hohen Ausprägungen von Komplexität deuten darauf hin, dass Leitstellen genau solche komplexen Umgebungen darstellen. In Bezug auf menschliches Verhalten in komplexen Umgebungen argumentiert Dörner (2008, S. 286–288), dass Menschen in solchen Systemen mit kognitiver Überlastung und suboptimalen Entscheidungsstrategien reagieren, wenn geeignete Aktionen fehlen, um gesetzte Ziele zu erreichen, womit die Ausführungen zur Komplexität in Leitstellen in Kapitel 2 empirisch gestützt werden. Bezüglich dieser Unterschiede beim geteilten mentalen Modell

kann auf Gräff et al. (2021) verwiesen werden, die bereits auf Unterschiede zwischen Rettungs- und Pflegefachpersonal im Schockraum hingewiesen haben.

Reflexion: Der Fragebogen fokussiert die subjektive Bedeutung aus Sicht der Teilnehmer:innen, anstatt Auftretenshäufigkeiten von Antwortoptionen zu messen. Ebenso können fragebogentypische Antworttendenzen (z. B. sozial erwünschte Antworten) nicht ausgeschlossen werden. Starre Trennwerte, wie sie im vorliegenden Kontext implementiert wurden, bergen das Risiko, dass subtile Differenzierungen übersehen werden und die Komplexität menschlicher Verhaltensweisen nicht adäquat erfasst wird.

Implikationen: Die Erkenntnisse zu Komplexität legen die Notwendigkeit von Maßnahmen zur Komplexitätsreduktion nahe. Eine Strategie richtet sich zunächst auf organisationaler Ebene an die Führungskräfte und Betreiber der Leitstellen, indem das bereits im Arbeitsschutz etablierte STOPV-Schema (Lüdeke, 2019, S. 1) auf die Leitstellen übertragen wird. Das STOPV-Schema folgt dem Grundgedanken, Gefahrenquellen im Rahmen des Minimierungsgebots zu substituieren und, wenn dies nicht möglich sein sollte, Maßnahmen zum sicheren Umgang mit der Gefahrenquelle zu etablieren. Zu den Grundpflichten des Unternehmers (hier des Leitstellenbetreibers) ist eine Bewertung der Gefährdungen zu zählen. Gefahren können für die Disponent:innen durch Überbelastung bestehen, jedoch auch für die unterschiedlichen Qualitätsdimensionen der Notrufbearbeitung. Zu den Gefährdungen gehören auch gefahrbringende Bedingungen, durch die eine Gefährdung wirksam werden kann, beispielsweise Umgebungsbedingungen, Zeitdruck, Arbeitsübergabe im Schichtbetrieb, Arbeitsprozesse (DGUV - Deutsche Gesellschaft für Unfallversicherung, 2020, S. 5) und Schichtdauer in der Leitstelle (Furangi et al., 2021). Abstrahiert man diesen Sachverhalt auf die limitierten menschlichen Fähigkeiten im Umgang mit Komplexität, kann Komplexität als eine gefährdungsbringende Bedingung verstanden werden, was die Anwendung des STOPV-Schemas nahelegt. Anlage 5 illustriert den Transfer des STOPV-Schemas auf Leitstellen. Die folgenden Ausführungen elaborieren mögliche Maßnahmen in absteigender STOPV-Hierarchie, wobei davon ausgegangen wird, dass eine Substitution bereits fruchtlos geprüft oder versucht wurde und die Ebene Substitution deswegen im Folgenden entfällt.

Hinblickend auf technische Maßnahmen eröffnen sich im Zuge der fortschreitenden Technologisierung vielfältige Möglichkeiten, um Verzögerungen in der Rückmeldung und Intransparenzen innerhalb gewisser Grenzen zu minimieren und somit auch der Komplexität entgegenzuwirken, z. B. Videoübertragungen mittels Drohnen zur Erzeugung eines frühen

Lagebildes (Sautter et al., 2023), zur Deichkontrolle bei Hochwasser oder zur Erkundung bei Gefahrgutunfällen (Kippnich, 2021), bei Großschadenslagen oder als Tauchdrohnen (Huf, 2023) sowie in der Waldbrandbekämpfung (Khan & Neustaedter, 2019). Die Freigabe der Mobiltelefon-Kameras durch die Anrufer:innen ist ein weiteres Beispiel zur Lageerkundung (Bolle et al., 2009; Sarma et al., 2017; Sýkora et al., 2022). Fairerweise muss bei der Diskussion des technisch Möglichen erwähnt werden, dass Disponent:innen hier in der Rolle des Empfängers von niederschwellig zu verarbeitenden Informationen und weniger in der Rolle des Bedieners weiterer anspruchsvoller Werkzeuge sein sollten, um nicht mit gut gemeinten Entlastungsangeboten eine Steigerung des Taskloads zu bewirken. Weiterhin ist auch zu beforschen, wie sich eine Videounterstützung bei der Notrufabfrage und Lageerkundung auf die psychische Gesundheit der Disponent:innen auswirken wird, wenn diese in hoher Frequenz mit eindrücklichen Bewegtbildern von mitunter dramatischen Szenen konfrontiert werden.

Die als problematisch wahrgenommenen Komplexitätsaspekte des Informationsflusses und der Eigendynamik unterstreichen die Relevanz von Verbesserungen der Interoperabilität, nicht nur zwischen Leitstellen und Rettungsmitteln, sondern auch unter den ca. 230 BOS-Leitstellen (Trautmann et al., 2022, S. 6) selbst sowie weiteren kritischen Infrastrukturen. Diese Interoperabilität ist gegenwärtig lediglich rudimentär ausgebaut. Zu ähnlichen Ergebnissen kommen auch die Expert:innen der sog. „Fürther Gespräche“, des Expertenforums für den Rettungsdienst. So werden beispielsweise synchronisierte Abfragesystematiken zwischen den BOS-Leitstellen (112) und dem ärztlichen Bereitschaftsdienst (116117) sowie ein gemeinsamer Ressourceneinsatz bei „minderschweren“ Einsätzen vorgeschlagen (Institut für notfallmedizinische Bildung, 2024, S. 7), woraus sich das Potenzial einer Verbesserung des Informationsflusses zwischen diesen beiden Instanzen ergibt.

Der den Berufsbildern inhärenten Eigendynamik kann im Bereich der Leitstellen weiterhin zumindest partiell durch den Einsatz von Entscheidungsunterstützungssystemen und die Automatisierung repetitiver Abläufe entgegengewirkt werden (Maletzki et al., 2023). Bei den intelligenten Entscheidungsunterstützungssystemen in Leitstellen hat jedoch ebenso zu gelten, dass die durch die Endanwender:innen und wahrgenommene Nützlichkeit und Benutzerfreundlichkeit die oberste Zielebene der Entwicklung und Implementierung sind. In der Entwicklung kann diesbezüglich die Anwendung der DIN EN ISO 9241-220 (Deutsches Institut für Normung e. V., 2019), die die Prozesse zur Ermöglichung, Durchführung und

Bewertung menschenzentrierter Gestaltung für interaktive Systeme fokussiert, nahegelegt werden. In vorgenannter Norm werden beispielsweise die Endanwender:innen in den gesamten iterativen Entwicklungsprozess einbezogen und das Design auf einem umfassenden Verständnis der Endanwender:innen, Aufgaben und Arbeitsumgebungen aufgebaut. Weiterhin sind über die Entwicklungsphase hinaus geeignete Implementierungsstrategien zu wählen, beispielsweise angelehnt an die ISO/IEC 12207 (Deutsches Institut für Normung e. V., 2017), die den kompletten Softwarelebenszyklus berücksichtigt. Eine Berücksichtigung dieser Prozesse im Rahmen einer formativen (iterativ während Entwicklung und Implementierung) und summativen (abschließend zum Projektende) Projektevaluation ist empfehlenswert, um den Erfolg und die Effektivität des Projekts anhand vorher festzulegender Kriterien, die sich hauptsächlich an menschenzentrierten Anforderungen ausrichten, zu bewerten. Dabei ist zu beachten, dass weder „Menschenzentrierung“ noch „Evaluation“ beliebig zu interpretierende Begriffe sind. Dies stellt Große (2020) für „Menschenzentrierung“ und die Deutsche Gesellschaft für Evaluation (2016) für „Evaluation“ deutlich heraus.

Hinsichtlich der Anwendung von Checklisten zur Komplexitätsreduktion, wie sie auch im Leitstellenkontext bereits mancherorts genutzt werden, erweist sich der „Read-Do-Ansatz“ als vielversprechend (Greig et al., 2023; St Pierre et al., 2017). Dieser Ansatz fokussiert das unmittelbare Lesen und Ausführen einzelner Handlungsanweisungen, im Gegensatz zum „Do-Confirm-Ansatz“, bei dem eine Sequenz von Handlungen ausgeführt und anschließend verifiziert wird. Der Read-Do-Ansatz scheint insbesondere für Notfallsituationen prädestiniert, da er die sequenzielle Abarbeitung von Aktionen in einer definierten Reihenfolge gewährleistet und somit das Risiko minimiert, dass in zeitkritischen Situationen essenzielle Schritte übergangen werden. Somit erscheint die Anwendung einer standardisierten Notrufabfrage, bei der „Read-Do“ bereits die Prämisse ist, auch aus dieser Perspektive angezeigt. Es liegt nahe, dass fehlende Standardisierung in der Notrufabfrage kaum zu einer Komplexitätsreduktion beitragen dürfte, weshalb eine Standardisierung in der Notrufabfrage essenziell ist.

Darüber hinaus sind komplexitätssteigernde Faktoren auch in der mangelnden Interoperabilität von Informationssystemen zu verorten, beispielsweise durch die Trennung des ELS von internetbasierten Diensten (beispielsweise Karten-/Routingdiensten). Dies manifestiert sich weiterhin in Situationen, in denen diverse Informationsquellen wie Funkmeldeempfänger, Navigationssysteme und Alarmfax manuell abgeglichen werden müssen, da ein automatisierter Datenaustausch nicht realisierbar ist.

In der Studie von Trautmann et al. (2022, S. 36) gaben 74,12 % (n = 63) der Leitstellen an, weitere, nicht originäre Aufgaben zu übernehmen, z.B. „Wachaufgaben im Kreishaus (Schließdienst)“ oder die „Ausgabe von Übernachtungszugängen für das Obdachlosenwohnheim des DRK“. Es ist anzunehmen, dass die Reduktion dieser völlig „fachfremden“ Aufgaben ebenso zu einer Absenkung des Komplexitätsniveaus führen dürfte. Die Erkenntnisse zu Kontextfaktoren implizieren, dass einerseits die individuelle Ebene und Teamebene zur Problemlösung zu adressieren sind, andererseits aber auch kulturelle Aspekte in der Organisation nicht übersehen werden dürfen. Auf individueller Ebene und Teamebene können entsprechend kompetenzorientierte Lernziele in Fallsimulationen adressiert werden, wobei Taylor et al. (2024) explizit dafür werben, die Lehreinrichtungen stärker in die Pflicht zu nehmen und die indirekten, subtilen Herausforderungen, die in interprofessionellen Teams bereits weit verbreitet sind, zu erkennen, wahrzunehmen und darauf zu reagieren. Fallsimulationen haben sich auch im Leitstellenkontext bereits bewährt (Meischke et al., 2017; Neidert, 2019). Auf Organisationsebene ist die Implementierung und kontinuierliche Verbesserung einer „Just Culture“ geeignet, eine offene Kommunikation zu fördern und potenzielle Sicherheitsprobleme zu erkennen (Finn et al., 2024). Eine „Just Culture“, also eine gerechte (Sicherheits-)Kultur, ist ein Umfeld, das einen offenen Dialog fördert, um sicherere Praktiken zu erleichtern. Dazu können auch routinemäßige Feedbackgespräche zwischen Disponent:innen und Führungskräften gehören. Vertiefte Hinweise zur Umsetzung der Just Culture sind beispielsweise bei Harvey und Sotardi (2017) zu finden. Außerdem kann auch im Zusammenhang mit der Leitstellenorganisation die Implementierung von Systemen zur Erfassung von Beinahezweischenfällen und Zwischenfällen, sog. Critical Incident Reporting Systems (CIRS), empfohlen werden, um nicht-adäquate und risikobehaftete Prozesse frühzeitig zu erkennen und ausbessern zu können.

Bezüglich der Adaptionserfordernisse ist zu berücksichtigen, dass es insbesondere in stressbehafteten und kritischen Konstellationen essenziell ist, dass ein Team sowohl sein Kommunikations- als auch sein Koordinationsverhalten anpasst und dabei auf bewährte Strategien zurückgreift, die zuvor in weniger anspruchsvollen, routinemäßigen Situationen, beispielsweise durch Fallbeispieltraining, internalisiert wurden. Ein erfolgreiches Modell aus der Praxis beschreiben Krugmann et al. (2021) am Beispiel der ILS München.

Diese Erkenntnisse zum geteilten mentalen Modell implizieren, dass sowohl die individuelle Ebene und Teamebene als auch kulturelle Aspekte innerhalb der Organisation adressiert werden müssen. Es ist davon auszugehen, dass gute Teamarbeit die Arbeitsbelastung

erheblich verringert und somit einen wichtigen Gesundheits- und Effizienzfaktor darstellt (Keunecke et al., 2019; Prottengeier et al., 2019). Es ist anzunehmen, dass das geteilte mentale Modell einer Patient:innenversorgung stark mit der Offenheit zur Kommunikation zusammenhängt (Reader et al., 2007). Im Praxistransfer bedeutet dies beispielsweise, dass sichere und effektive Kommunikationsmechanismen, wie sie auch von Pruitt und Liebelt (2010) konkretisiert werden, habitualisiert und eingefordert werden müssen. Geschlossene Gesprächskreisläufe und Übergabeschemata sind hier beispielhaft zu nennen, auch für die Leitstelle.

Ausblick: Die weitere Forschung sollte darin bestehen, größere Teilstichproben der Berufsgruppen zu erhalten, um Technikbereitschaft und Team-Arbeit-Kontext besser repräsentieren zu können. Weiterhin sind die Gründe für geringe Technikbereitschaft im Leitstellensetting weiter zu explorieren, günstigstenfalls mit qualitativer Forschung. Die Auswirkungen leitstelleninterner Prozesse, wie beispielsweise der Struktur der Notrufabfrage, auf die Komplexität und mentale Arbeitsbelastung sind bislang weitgehend unerforscht und bedürfen einer intensivierten wissenschaftlichen Betrachtung.

5.1.2 RQ 2 - Überzeugungssystem der Leitstellendisponent:innen

Einleitung: Diese theoriegenerierende qualitative Studie, die auf semi-strukturierten Interviews von Disponent:innen und weiteren Expert:innen beruht, wurde per Grounded-Theory-Methodologie analysiert. Es wurden acht Interviews mit insgesamt 31 Personen geführt, die insgesamt 619 Minuten dauerten. Hinsichtlich der Forschungsfrage ist anzumerken, dass belastbare Daten zu mentalen Modellen von Disponent:innen in Deutschland bisher fehlen. Auch wenn konkrete Ergebnisrepräsentation in der Diskussion eher unüblich sind, werden im Folgenden bewusst In-vivo-Zitate angeführt, um die Interpretation datengetrieben zu untermauern.

Interpretation: Die Interviews zeigen, dass die Arbeitsbelastung der Disponent:innen durchgängig als hoch angenommen wird. Auch wenn sich ein belastbarer Widerspruch des bereits dargestellten Zitats von Hofinger (S. 9) zu den Interviewdaten nicht einwandfrei darstellen lässt, kann eine erhebliche Entlastung der Disponent:innen durch die derzeitige Umsetzung von Standardprozeduren, Algorithmen oder Leitlinien jedoch bezweifelt werden, wie u. a. das folgende In-vivo-Zitat andeutet.

In-vivo-Zitat: „Wir sind eine Krake mit 10 Armen.“ (Fokusgruppen-Interview, 20.09.21, Person 6, Absatz 89)

Bezüglich der Einstellung der Disponent:innen zu KI konnte dargestellt werden, dass nicht nur wohlwollende und positive Überzeugungssysteme gegenüber KI-Systemen in Leitstellen vorhanden sind, sondern auch ablehnende Haltungen, die sich mit einer Betonung der menschlichen Leistungsfähigkeit verbinden.

In-vivo-Zitate: *„Also ich sage immer ich komme auch ohne System zurecht [...]“* (Expert:innen Interview 03.11.21, Absatz 77). *„Ja, ich glaube, eine KI sollte nicht mit dem Notfallpatienten reden.“* (Expert:innen Interview 10.10.23, Absatz 60). *„Es ist ja klar, dass die künstliche Intelligenz den Menschen einfach nicht stechen kann und es wird auch niemals so weit kommen.“* (Fokusgruppen-Interview, 20.09.21, Person 1, Absatz 45)

Es ist aufgrund der Antwortmuster davon auszugehen, dass beruflicher Status, KI-Wissen und Qualifikationen als positive Prädiktoren wirken. Geringe Ausprägungen führen zu Skepsis und einer „Blackbox“-Sicht auf KI. In dieser Studie bedeutet „Blackbox“, dass bestenfalls eine diffuse, nebulöse, volatile und manchmal sogar dystopische Sicht auf KI-Methoden besteht. Reflexionsfähigkeiten der Disponent:innen über den Arbeitskontext können jedoch zu positiveren Einstellungen und zu einer „Lichtung“ der Blackbox führen.

Teilweise sind die Beschreibungen der eigenen Arbeitskontexte durch die Leitstellendisponent:innen auch mit konkreteren Vorstellungen verbunden:

In-vivo-Zitate: *„Wir benötigen Unterstützung bei Automatisierung, aber auch bei der Visualisierung.“* (Fokusgruppen-Interview, 20.09.21, Person 6, Absatz 89). *„Ich hätte beispielsweise gerne zwei Monitore, so dass es einfach übersichtlicher und tätigkeitsbezogener funktioniert. [...] Da hat man so viele verschiedene Fenster auf, die man nur einmal am Tag oder gar nicht braucht. Das ist unnötig. Das kann man alles reduzieren, meiner Meinung nach. Auf mindestens zwei Monitore, wenn nicht sogar auf einen [Monitor].“* (Expert:innen-Interview 20.10.23, Absatz 19)

Die Erkenntnisse sind kongruent zu denen aus der Studie zur Technikbereitschaft (P1), indem sie ein Bild zeichnen, das Leitstellendisponent:innen grundsätzlich technikbereit darstellt, aber andererseits zeigt, dass es eben auch Ängste, Vorbehalte und Wissensdefizite gegenüber dem konkreten Einsatz von KI-Systemen gibt.

In-vivo-Zitate: *„Es geht auch darum, [...] alle haben Angst um Ihre Arbeitsplätze [...]“* (Expert:innen Interview 20.10.23, Absatz 19). *„Ich würde es bedenklich finden, wenn der Kontakt zu Anrufern durch eine KI ersetzt wird. Speziell im Sinne eines Notrufs, an anderen Stellen kann ich mir das sehr gut vorstellen.“* (Fokusgruppen-Interview, 20.09.21, Person 6, Absatz 43). *„Ich möchte [...] kein Programm, auf das ich aufpassen muss, sondern ein Programm, das auf mich aufpasst.“* (Fokusgruppen-Interview, 20.09.21, Person 1, Absatz 45)

Einordnung in den Forschungskontext: Hinsichtlich des Bezugs zu einer übergeordneten Theorie kann hier auf Festingers Theorie der kognitiven Dissonanz verwiesen werden. Diese besagt, dass Menschen ein starkes Bedürfnis nach Konsistenz zwischen ihren Überzeugungen,

Einstellungen und Verhaltensweisen haben. Wenn Inkonsistenzen auftreten, entsteht ein unangenehmer psychologischer Zustand (kognitive Dissonanz), den Individuen zu reduzieren versuchen, indem sie ihre Überzeugungen ändern, neue Informationen suchen oder die Bedeutung der dissonanten Elemente herunterspielen (Cooper, 2007; Festinger, 1957). Im Kontext der Studie kann dies erklären, wie Disponent:innen mit widersprüchlichen Gefühlen gegenüber KI umgehen, z. B. zwischen der Anerkennung potenzieller Vorteile und Bedenken hinsichtlich der Jobsicherheit oder Komplexität, wie es auch in den Interviews gezeigt werden konnte. Besonders die Hinweise auf die menschliche Leistungsfähigkeit gegenüber den Fähigkeiten der KI können als dysfunktionaler Lösungsversuch der Dissonanzreduktion verstanden werden.

Reflexion: Bezüglich der Limitation der Studie ist zu erwähnen, dass nationale Stichprobe und Kontexte nicht ohne Weiteres auf internationale Fragestellungen übertragbar sind. Weiterhin fokussiert sich die Studie auf KI-basierte Entscheidungsunterstützungssysteme in Leitstellen und eine Verallgemeinerung auf andere KI-Kontexte erfordert weitere Überlegungen. Auch ist zu berücksichtigen, dass eine Standardisierung des Begriffs „KI“ in dieser Studie bewusst nicht erfolgt ist und die interviewten Personen heterogene Vorstellungen von „der KI“ zeigen. In Bezug auf die Qualitätskriterien nach Steinke (2019, S. 319–331) können intersubjektive Nachvollziehbarkeit, die Indikation der gewählten Methoden, empirische Verankerung der Theorie in den Daten, Kohärenz und Relevanz angenommen werden. Die reflektierte Subjektivität wurde durch umfangreiche Diskussionen der Methoden und Ergebnisse mit einem nicht an der Forschung beteiligten Experten verbessert. Da nur Personen befragt wurden, die noch in ihrem Beruf aktiv sind, was ein gewisses berufliches Engagement impliziert, und nicht Berufsausstieger, kann es auch zu einer Verzerrung im Sinne eines Survivorship-Bias kommen.

Implikationen: Diese Ergebnisse verdeutlichen den hohen Stellenwert einer partizipativen Entwicklung, Integration und Information (Dissemination), die sich an den tatsächlichen Bedürfnissen der Leitstellendisponent:innen sowie den organisatorischen und taktischen Anforderungen ausrichten, was gleichermaßen die Ausbildung betrifft. Dies ist von erheblichem Stellenwert, um dem Vorwurf der praxisfernen Wissenschaft im Elfenbeinturm zu entgehen, wie es auch Westerheide und Schott (2022) im Kontext der Humanisierung des Arbeitslebens darstellen. Hinsichtlich der Entwicklung und Integration kann als Lösungsansatz auf das Konzept der Explainable Artificial Intelligence (XAI) verwiesen werden, neben der

konsequenter Anwendung nutzerzentrierter Methoden, wie bereits weiter oben beispielsweise durch Anwendung der DIN EN ISO 9241-220 (siehe Kapitel 5.1.1.3) vorgeschlagen. Es werden drei Hauptbestandteile der XAI unterschieden (Minh et al., 2022), die sich auch auf den Leitstellenkontext übertragen lassen: Erklärbarkeit vor der Modellierung, interpretierbare Modelle und Erklärbarkeit nach der Modellierung. Umsetzungsimpulse werden in Anlage 6 expliziert. Weiterhin kann aufgrund deren Blackbox-Charakteristik diskutiert werden, wie geeignet sehr eingeschränkt erklärbare KI-Verfahren, wie z. B. künstliche neuronale Netzwerke, für DSS in der Leitstelle sind und ob nicht grundsätzlich gut erklärbare Methoden der mathematischen Optimierung bevorzugt angewendet werden sollten, solange nicht zwingende Gründe diese ausschließen.

Ausblick: Die vorgeschlagenen Maßnahmen der Entwicklung, Integration, Information und Ausbildung sind auf wissenschaftliche Art und Weise gemeinsam mit den fachpraktischen Akteur:innen umzusetzen und zu evaluieren. Maßnahmen der Komplexitätsreduktion erscheinen besonders dringlich.

5.1.3 RQ 3 - Aufmerksamkeitsverteilung der Disponent:innen beim Notrufdialog

Einleitung: In einer Vorarbeit (P5) wurden iterative Schritte zur grundsätzlichen Implementierung und Evaluation von hybrid-intelligenten DSS entwickelt. Jedoch fehlen konkrete Erkenntnisse zu den Arbeitsprozessen der Disponent:innen, um Visualisierungskonzepte innerhalb eines DSS entwickeln zu können. Deshalb wurde in P6 zwei ILS mit unterschiedlichen technischen Set-ups - im Wesentlichen zwei Hauptmonitore und vier Hauptmonitore eine Eyetracking-Studie durchgeführt, mit dem Ziel, die visuelle Aufmerksamkeit zur Ableitung von Designvorschlägen für ein DSS zu erheben. Zusammenfassend konnte die Eyetracking-Studie auf eine Stichprobengröße von $N = 33$ zugreifen.

Interpretation: Die entwickelten Schritte der Implementierung und Evaluation von DSS konnten sich bereits in der Konfrontation mit Expert:innen aus dem Leitstellenbereich bewähren (P5).

Bezüglich der Stichprobengröße von P6 ($N = 33$) ist zu berücksichtigen, dass es keine einheitlichen Empfehlungen für Eyetracking-Studien gibt, wie auch Sharafi et al. (2020, S. 19) berichten. Bei Eyetracking-Studien ist mit 5 bis 169 Teilnehmer:innen eine große Bandbreite zu verzeichnen (Sharafi et al., 2020 S. 19). Die Eyetracking-Analyse konnte durch die Berechnung der Fixation pro AOI, die dem Bildschirm entspricht, zeigen, dass vor allem im

Vier-Monitor-Set-up den beiden äußeren Monitoren wesentlich weniger visuelle Aufmerksamkeit zugewendet wird als den beiden zentralen Monitoren. Weiterhin konnte durch das Adressieren des Raw-TLX-Fragebogens an die Teilnehmer:innen gezeigt werden, dass in Leitstelle 1 signifikant höhere Niveaus von mentaler Belastung und Zeitdruck erfahren wurden. Dies wirft Fragen zur Hard- und Softwareergonomie des Arbeitsplatzes auf und ist kritisch zu diskutieren, auch im Hinblick auf konkurrierende Reizwahrnehmung mit Hinweis auf die bereits erwähnte exogene visuelle Aufmerksamkeitsselektion (siehe Kapitel 3.2.4). Eine weitere Begründung liefert die der Lehr-Lern-Theorie entlehnte Cognitive-Load-Theorie (CLT) nach Sweller (1988), wie die Einordnung in der Forschungskontext zeigen wird.

In Bezug auf die angesprochene Ergonomie kann eine Reduktion der Monitore in Verbindung mit einer kontextsensitiven Visualisierung, also der Visualisierung der tatsächlich benötigten Information, und einer Unterdrückung ablenkender visuelle Stimuli, die nicht zur Zielerreichung beitragen, hilfreich sein. Diese Hypothese gilt es in Folgeforschungen zu prüfen.

Einordnung in den Forschungskontext: Grundlage für die erarbeiteten Empfehlungen in P6 zur Evaluation sind die Evaluationsstandards der Deutschen Gesellschaft für Evaluation (2016). In Bezug auf die Mensch-KI-Interaktion demonstriert der Transfer der Leitlinien für Mensch-KI-Interaktion (Amershi et al., 2019) auf die Leitstellen die Verbindung mit der aktuellen Forschung (Amershi et al., 2019). Hinsichtlich hybrid-intelligenter Systeme wurden relevante Studien im Sozio-Informatik-Bereich berücksichtigt, beispielsweise von Akata et al. (2020).

Bezüglich der Einordnung von P6, ist auf die CLT hinzuweisen. Diese ist eine einflussreiche Theorie, die sich mit der kognitiven Belastung beim Lernen befasst (Schnotz & Kürschner, 2007, S. 470). Sie geht davon aus, dass das Arbeitsgedächtnis in seiner Kapazität begrenzt ist und dass Lernen durch eine Überlastung des Arbeitsgedächtnisses behindert werden kann. Da sich die CLT zwar auf Instruktionsdesign bezieht, aber weitläufige und elaborierte Annahmen zur menschlichen Wahrnehmungsarchitektur („human cognitive architecture“) beinhaltet (Sweller, 2011), lässt sich die Theorie auch auf andere Bereiche, wie die Leitstelle, anwenden. Die CLT unterscheidet zwischen intrinsischer und extrinsischer kognitiver Belastung (Sweller, 2011, S. 57–72). Intrinsische Belastung bezieht sich auf die inhärente Komplexität der zu verarbeitenden Sachverhalte und wird hauptsächlich durch die Anzahl der Informationselemente, die gleichzeitig im Arbeitsgedächtnis verarbeitet werden müssen, bestimmt. Die intrinsische Belastung ist durch die Aufgabenstellung an sich determiniert.

Bezogen auf die Arbeitsrealität der Disponent:innen kann davon ausgegangen werden, dass ein erheblicher Teil der intrinsischen Belastung durch die sehr unterschiedlichen Einsatzlagen vor Ort entstehen, die nicht veränderlich sind, da sich die Situation (z. B. Verkehrsunfall) bereits verwirklicht hat.

Extrinsische Belastung hingegen resultiert aus der Art und Weise, wie Informationen präsentiert werden, und nicht aus der inhärenten Komplexität der Lage selbst. Ein Beispiel hierfür ist die Belastung, die durch die Interaktion mit dem ELS und anderer Soft- und Hardware entsteht. Im Gegensatz zur intrinsischen Belastung kann die extrinsische Belastung durch Veränderungen an der Mensch-Maschine-Schnittstelle beeinflusst werden, ohne das taktische Ziel zu verändern.

Die bisherigen Überlegungen sind jedoch auch im Kontext des Eyetrackings zu verstehen, weshalb die Eye-Mind-Hypothese (EMH) (Just & Carpenter, 1980) zu beobachten ist. Diese nimmt an, dass die Augen das Objekt fixieren, das gerade im Fokus der Aufmerksamkeit ist und geht somit von einer starken Verbindung zwischen Blickfixationen und mentalen Prozessen aus. Neuere Arbeiten betonen, dass der Zusammenhang zwischen Blickbewegungen und kognitiven Prozessen von Aufgaben- und Domäneneigenschaften abhängt und dass Fixationen nicht unbedingt die aktuellen kognitiven Prozesse widerspiegeln (Schindler & Lilienthal, 2019), sondern auch vergangene oder zukünftige Informationsverarbeitung (Wu & Liu, 2022) und von Vorwissen abhängen können (Epelboim & Suppes, 2001). Weiterhin ist davon auszugehen, dass bei Aufgaben mit größerem Interpretationsspielraum die EMH weniger wahrscheinlich zutrifft (Schindler & Lilienthal, 2019). Trotz dieser metatheoretischen Limitationen kann von einer grundlegenden Validität der EMH im Kontext dieser Studie ausgegangen werden, besonders da in Leitstellen die visuelle Informationsverarbeitung durch die Gestaltung der grafischen Bedienoberflächen zentral ist. Weiterhin, so argumentieren auch Wu und Liu (2022), kann die EMH bei spezifischen, klar definierten Teilaufgaben in der Leitstelle (z. B. der Notrufabfrage, die hier adressiert wird) durchaus zutreffen. Abgesehen davon bietet Eyetracking eine hohe zeitliche Auflösung, die andere Methoden oft nicht erreichen, was besonders in zeitkritischen Leitstellensituationen wertvoll ist.

Reflexion: Zu P5 ist zu erwähnen, dass zwar eine positive Validierung der Ergebnisse mit Expert:innen aus dem Feld erfolgte, was vor dem spezifischen Projekthintergrund als angemessen beurteilt werden kann. Da die Expert:innen lediglich aus Rheinland-Pfalz

rekrutiert wurden, ist die Generalisierbarkeit der Ergebnisse nur eingeschränkt möglich. Hinsichtlich der Elemente der Evaluation ist jedoch auf die bereits dargestellten hohen Leitlinienadhärenz hinzuweisen.

Zu berücksichtigen ist bei P6, dass bei dieser Feldstudie eine Situation *in situ* beforscht wird und sich in einer komplexen Leitstellenumgebung die Kontrolle der internen Validität anspruchsvoll gestaltet. So können Konfundierungseffekte, beispielsweise durch unterschiedliche organisatorische Prozesse, wie Sitzanordnung, die eine abweichende Kommunikation zwischen Calltaker und Dispatcher erfordern, die kausale Zuschreibung von Erklärungen auf die visuelle Aufmerksamkeit und Dimension des Raw-TLX erschweren. In Folgestudien müssen dezidierte Maßnahmen zur Drittvariablenkontrolle diskutiert werden.

Die Schwankungen in der Datenqualität sind in Eyetracking-Studien nicht ungewöhnlich. Schnipke und Todd (2000), Mullin et al. (Mullin et al., 2001) und Pernice und Nielsen (2009, nach Holmqvist et al., 2012) berichten von Datenverlusten von 20 - 60 % der Teilnehmer:innen und Burmester und Mast (2010) schlossen 12 von 32 Teilnehmer:innen aufgrund von Problemen mit der Datenqualität oder der Kalibrierung aus. In der hier vorliegenden Studie sind die Datenverluste bei Brillenträger:innen auf Störreflexionen zurückzuführen, die ein Tracking der Pupillen erschweren und nachfolgend zu fehlenden Werten führen. In anderen Fällen kann ein unbeabsichtigtes und unbemerktes Verrutschen des Eyetrackers, der sog. Drift (Sharafi et al., 2020, S. 7), vermutet werden. Hier ist bei zukünftigen Eyetracking-Studien zu überprüfen, ob kürzere Aufnahmezeiten zu weniger Drift und einer höheren Datenqualität führen. Weiterhin wäre kritisch zu prüfen, ob eine eigens dafür abgestellte wissenschaftliche Hilfsperson, das Tracking während des laufenden Versuchs in Echtzeit überprüfen sollte, um eine neue Kalibration anzustoßen. Dies würde jedoch eine Nutzen-Risiko-Abwägung mit Praxistests benötigen, da bei der genutzten Kalibrationsmethode bei jeder Kalibration die aktive Mitwirkung der Versuchsperson benötigt wird (Santini et al., 2017) und diese unweigerlich zu einem Präsenzbruch führt, die Versuchsperson also aus ihrer Aufgabe „herausgeholt“ wird.

Zu den fehlenden Werten sind allerdings auch physiologische Überlegungen zu berücksichtigen. In der Analyse von Blickbewegungsdaten ist es ein normatives Phänomen, dass Blickpunkte existieren, die nicht als Fixationen klassifiziert werden. Dies ist ein inhärenter Bestandteil der okulomotorischen Dynamik. Bei einwandfreier Funktionalität der Pupillendetektion lässt sich aus den erfassten Daten stets ein Blickpunkt berechnen. Während sakkadischer Bewegungen, also schneller bewusster Augenbewegungen, findet in der Regel

keine bewusste Wahrnehmung statt (sakkadische Suppression). Folglich ist es methodologisch korrekt und erwartungsgemäß, dass während dieser Phasen keine spezifischen AOIs registriert oder ausgegeben werden.

Implikationen: In Bezug auf die Leitstelle bedeutet dies, dass die Art und Weise, wie Daten auf den verschiedenen Bildschirmen zu Informationen zusammengefasst und visualisiert werden, Einfluss auf die extrinsische kognitive Belastung nehmen könnte und an Designüberlegungen auszurichten ist, die am Stat-of-the-Art, beispielsweise den Leitlinien zu Mensch-KI-Interaktion (Amershi et al., 2019), orientiert werden sollten. Leitliniengerechte Evaluationen der Systeme in Hinblick auf die Mensch-KI-Interaktion sind notwendig. Werden von vier Monitoren zuzüglich Touchscreen hauptsächlich nur zwei Monitore genutzt, stellt sich nicht nur die Frage der Notwendigkeit von vier versus zwei Monitoren, sondern auch die Frage, ob durch die zusätzlichen und mutmaßlich nicht notwendigen Monitore die extrinsische Belastung gesteigert wird. Diese Hypothese kann durch die Annahme gestützt werden, dass die beiden äußeren Monitore eine gewisse Salienz besitzen. Besonders wenn diese noch Farbwechsel anzeigen, beispielsweise wenn sich der Status eines Rettungsmittels ändert, ist diese Salienz nochmals höher. Erklärt werden kann dies mit dem konkurrierenden Verhältnis von reiz- und zielinduzierter Aufmerksamkeit (Gerrig, 2018, S. 158–159). Während es bei zielinduzierter Aufmerksamkeit darum geht, welche Objekte aufgrund eigener Ziele fokussiert werden, beschreibt die reizinduzierte Variante die Aufmerksamkeitslenkung unabhängig von eigenen Zielen, die in der Situation priorisiert werden sollten, wie Gerrig ibidem ausführt. Als Ableitung für die Leitstellenpraxis kann deshalb gelten, die Notwendigkeit konkurrierender Reize in die Planung von Soft- und Hardware miteinzubeziehen und Information zukünftig „intelligent“, also kontextsensitiv, zu visualisieren. Dabei können DSS, die eine intelligente Visualisierung von angemessenen Informationsmengen (Manzi et al., 2018) mit intuitiven Symbolen beinhalten, kontextsensitives Feedback (Wiczorek & Meyer, 2019) geben und an die sehr heterogenen Leitstellenarbeitsplätze anpassbar sind, hilfreich sein. Dass Disponent:innen dabei zuerst nach passenden Handlungsanweisungen innerhalb der IT-Systeme suchen müssen, wie es in Kapitel 2 Komplexität in Leitstellen skizziert wird, erscheint nicht hilfreich.

Ausblick: Auch wenn die Stichprobengröße in Eyetracking-Studien nicht unüblich ist, sind weitere Studien durchzuführen, um den variablen technischen Set-ups der Leitstelle begegnen zu können, um durch Standardisierung der Aufgaben eine bessere Vergleichbarkeit erreichen

zu können und um letztendlich Drittvariablen besser kontrollieren zu können. Die Wirkung von kontextsensitiven Visualisierungen und Feedback des DSS auf Parameter wie die aufgabenbezogene mentale Arbeitsbelastung und die visuelle Aufmerksamkeitsverteilung von Disponent:innen ist zu beforschen. Feldstudien zu Designanforderungen von KI-Anwendungen in den Leitstellen sind auszuweiten.

5.2 Ausblick und weitere Forschung

Zukünftige Forschung sollte auf die Erweiterung der Teilstichproben verschiedener Berufsgruppen abzielen, um Technikbereitschaft und Team-Arbeit-Kontext präziser zu erfassen. Eine tiefergehende Exploration der Ursachen geringer Technikbereitschaft im Leitstellenumfeld, vorzugsweise durch qualitative Ansätze, ist erforderlich. Die dadurch generierten Erkenntnisse sind für die zukünftige Gestaltung von Personalauswahl und Qualifizierungsmaßnahmen zu berücksichtigen. Zudem bedürfen die Auswirkungen leitstelleninterner Prozesse auf Komplexität sowie mentale Arbeitsbelastung und daraus folgend auf Versorgungsqualität und Patient:innensicherheit einer intensiveren wissenschaftlichen Betrachtung.

Im Bereich menschenzentrierter Designs ist eine Ausweitung der Forschung direkt in den Leitstellen anzustreben, um Arbeitskontexte, -abläufe und -bedürfnisse der Disponent:innen besser zu verstehen. Eine Erweiterung der Eyetracking-Feldstudie mit einer größeren Stichprobe, einem Vergleich verschiedener Monitorkonfigurationen und standardisierten Aufgaben in Form einer Leitstellensimulation wäre wünschenswert. Die Ergänzung des Selbstberichtsverfahrens zur Taskload-Erfassung durch physiologische Messungen könnte zur Objektivierung beitragen.

Die vorgeschlagenen Maßnahmen sollten wissenschaftlich fundiert und in Kooperation mit fachpraktischen Akteur:innen umgesetzt und evaluiert werden. Besonders dringlich erscheint die Komplexitätsreduktion durch die partizipative Entwicklung und Implementierung menschengerechter Entscheidungsunterstützungssysteme, die sowohl den Bedürfnissen der Disponent:innen als auch modernen Designkriterien entsprechen.

Ein zukünftiges Anliegen kann auch unter der Prämisse der Salutogenese („Gesunderhaltung“) von Disponent:innen verstanden werden, da sowohl komplexitätsbringende als auch resilienzfördernde Faktoren herausgearbeitet werden können. Vor dem Hintergrund der Burn-out-Epidemie (Alkærsig et al., 2018) in Organisationen, die auch die Retter:innen betrifft (Baier et al., 2018), scheint dieses Anliegen geradezu die Qualität eines kategorischen

Imperativs anzunehmen, auch weil bei Disponent:innen im Rahmen einer möglichen Second-Victim-Problematik zusätzliche Belastungen anzunehmen sind (Marung et al., 2023).

5.3 Fazit

Die vorliegende Dissertation widmete sich der Entwicklung und Evaluation intelligenter Systeme für Leitstellen der Behörden und Organisationen mit Sicherheitsaufgaben (BOS) unter besonderer Berücksichtigung der Perspektive der Endanwender:innen. Diese Analyse zielt darauf ab, eine datengetriebene Grundlage für die Konzeption von Unterstützungssystemen, insbesondere KI-Verfahren, zu schaffen, die den spezifischen Anforderungen und Bedürfnissen der Leitstellendisponent:innen gerecht werden und daraus resultierend auch durch eine stabile Nutzungstreue positiven Einfluss auf Patient:innensicherheit und Versorgungsqualität nehmen.

Zum Fazit von **Forschungsfrage 1** hinsichtlich der **Demografie** kann festgehalten werden, dass der Beruf „Leitstellendisponent:in“ hauptsächlich durch berufs- und lebenserfahrene Männer ausgeübt wird. Daraus ergibt sich einerseits die Notwendigkeit, die Erfahrungen und Biografien auch in Qualifikationsmaßnahmen hinsichtlich neuer Technologien in Leitstellen spezifisch zu berücksichtigen, aber auch, jüngere Menschen, die vor allem im jugendlichen und jungen Erwachsenenalter eine hohe Aufmerksamkeit auf soziale Medien verwenden, nicht aus dem Blick zu verlieren, was sowohl die Rekrutierung als auch die Bildung betreffen kann. Ob die Altersverteilung bei der Einführung eines dreijährigen Ausbildungsberufs „Leitstellendisponent:in“ auf diesem Niveau bleiben wird, ist fraglich.

In puncto **Technikbereitschaft** sind zwar hauptsächlich positive Ausprägungen zu beobachten, jedoch auch wenige Menschen mit sehr geringer Technikbereitschaft, was die Notwendigkeit gezielter Maßnahmen notwendig macht, beispielsweise Kommunikationsstrategien zur frühen Darstellung von beruflicher Relevanz und Nützlichkeit. Weiterhin sind „KI-Kompetenzen“ zielgerichtet, d. h. auch rollenspezifisch, in den Leitstellen zu vermitteln und neue technische Durchstiche möglichst früh an die Disponent:innen zu kommunizieren, um Fürsprecher:innen und Multiplikator:innen zu gewinnen.

Bei der Beforschung des **Team-Arbeit-Kontexts** hat sich die Notwendigkeit von komplexitätsreduzierenden Maßnahmen für die Disponent:innen gezeigt. Der Transfer des aus dem Arbeitsschutz bekannten STOPV-Schemas auf die Komplexität im jeweiligen individuellen Leitstellenkontext erscheint naheliegend, denn Komplexität ist als gefahrbringende Bedingung zu verstehen. Weiterhin sind technische Möglichkeiten zu

berücksichtigen, die Informationsdefizite und Intransparenzen reduzieren, indem sie komplexe Datenströme zu einfach verstehbaren Informationen zusammenführen, die im Entscheidungsunterstützungssystem direkt verarbeitet werden. Dazu gehören beispielsweise Drohnenbilder der Lage vor Ort oder Videos, Bilder und weitere Daten der Mobiltelefone von Beteiligten und Passant:innen.

Zieht man zur Gestaltung der **Überzeugungssysteme** aus **Forschungsfrage 2** ein Fazit, kann der hohe Stellenwert einer partizipativen Entwicklung, Integration und Information (Dissemination) betont werden. Diesbezüglich sind die Leitstellendisponent:innen als zentrale Stakeholdergruppe von KI-Systemen in Leitstellen hervorzuheben, die es entsprechend zu berücksichtigen gilt, um negative oder gar aversive Haltungen gegenüber „der KI“ zu vermeiden. Weiterhin ist neben dem bereits angesprochenen Adressieren von KI-Kompetenzen auf die Anwendung bereits vorhandener nutzerzentrierter Standards als absolute Notwendigkeit hinzuweisen. Einer erklärbaren und verstehbaren KI das Primat einzuräumen, scheint in diesem Bezug unhintergebar.

Als Fazit zu **Forschungsfrage 3** in Bezug auf die Verteilung der **visuellen Aufmerksamkeit** der Disponent:innen ist auf deutliche Hinweise abzuheben, die annehmen lassen, dass mehr Monitore am Leitstellenarbeitsplatz nicht zu einer geringeren aufgabenspezifischen mentalen Arbeitsbelastung führen. Eine kontextsensitive Visualisierung von Informationen, eingebettet in ein intelligentes Entscheidungsunterstützungssystem, scheint einen höheren Stellenwert in Sachen Visualisierung zu haben. Durch die Beschreibung von wenigen Schritten der Implementierung und Evaluation von hybrider KI konnte außerdem gezeigt werden, dass sowohl die Implementierung von hybrider KI als auch die Evaluation im Leitstellensetting wissenschaftlich und zeitgleich praxisnah möglich sind. Die Evaluation ist dabei mit hoher Leitlinienadhärenz durchzuführen, um eine angemessene Ergebnisqualität zu erzielen.

In **zukünftiger Forschung** sollten vor allem weitere Subgruppenanalysen durch noch größere Stichproben ermöglicht, der Heterogenität der Leitstellensettings durch weitere Feldforschung verstärkte Beachtung geschenkt und in Feldforschungen Konzepte zur Reduktion der Wirkung von möglichen Drittvariablen berücksichtigt werden.

Literaturverzeichnis

- Akata, Z., Balliet, D., Rijke, M. de, Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F. & Welling, M. (2020). A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. *Computer*, 53(8), 18–28. <https://doi.org/10.1109/MC.2020.2996587>
- Alkærsg, L., Kensbock, J. & Lomberg, C. (2018). The Burnout Epidemic—How Burnout Spreads Across Organizations. *Academy of Management Proceedings*, 2018(1), 14180. <https://doi.org/10.5465/AMBPP.2018.14180abstract>
- Allport, D. A. (1987). Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Hrsg.), *Perspectives on Perception and Action* (S. 395–419). Lawrence Erlbaum Associates.
- Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R. & Horvitz, E. (2019). Guidelines for Human-AI Interaction. In S. Brewster, G. Fitzpatrick, A. Cox & V. Kostakos (Hrsg.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (S. 1–13). ACM. <https://doi.org/10.1145/3290605.3300233>
- Aviz, I. L., Souza, K. E., Ribeiro, E., Mello Junior, H. de & Da Seruffo, M. C. R. (2019). Comparative study of user experience evaluation techniques based on mouse and gaze tracking. In J. dos Santos, D. C. Muchaluat Saade, M. Da Graça C. Pimentel & A. A. Macedo (Hrsg.), *Proceedings of the 25th Brazillian Symposium on Multimedia and the Web* (S. 53–56). ACM. <https://doi.org/10.1145/3323503.3360623>
- Baier, N., Roth, K., Felgner, S. & Henschke, C. (2018). Burnout and safety outcomes - a cross-sectional nationwide survey of EMS-workers in Germany. *BMC emergency medicine*, 18(1), 24. <https://doi.org/10.1186/s12873-018-0177-2>
- Baubin, M., Häske, D., Lechleuthner, A. & Luiz, T. (2020). Fokus Leitstelle – „Alarmierst Du nur oder berätst Du schon?“ [Dispatch center focus issue-do you just sound the alarm or do you already advise?]. *Notfall & rettungsmedizin*, 23(7), 487–489. <https://doi.org/10.1007/s10049-020-00796-3>
- Bednarik, R., Eivazi, S. & Hradis, M. (2012). Gaze and conversational engagement in multiparty video conversation. In *Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction* (S. 1–6). ACM. <https://doi.org/10.1145/2401836.2401846>
- Berger, W., Coutinho, E. S. F., Figueira, I., Marques-Portella, C., Luz, M. P., Neylan, T. C., Marmar, C. R. & Mendlowicz, M. V. (2012). Rescuers at risk: a systematic review and meta-regression analysis of the worldwide current prevalence and correlates of PTSD in rescue workers. *Social psychiatry and psychiatric epidemiology*, 47(6), 1001–1011. <https://doi.org/10.1007/s00127-011-0408-2>
- Blomberg, S. N., Christensen, H. C., Lippert, F., Ersbøll, A. K., Torp-Petersen, C., Sayre, M. R., Kudenchuk, P. J. & Folke, F. (2021). Effect of Machine Learning on Dispatcher Recognition of Out-of-Hospital Cardiac Arrest During Calls to Emergency Medical Services: A Randomized Clinical Trial. *JAMA network open*, 4(1), e2032320. <https://doi.org/10.1001/jamanetworkopen.2020.32320>
- Blomberg, S. N., Folke, F., Ersbøll, A. K., Christensen, H. C., Torp-Pedersen, C., Sayre, M. R., Counts, C. R. & Lippert, F. K. (2019). Machine learning as a supportive tool to recognize

- cardiac arrest in emergency calls. *Resuscitation*, 138, 322–329. <https://doi.org/10.1016/j.resuscitation.2019.01.015>
- Bogner, A., Littig, B. & Menz, W. (2014). *Interviews mit Experten: Eine praxisorientierte Einführung*. Springer VS.
- Bundeszentrale für politische Bildung. (2024). *SINUS-Jugendstudie 2024 – „Wie ticken Jugendliche?“*. <https://www.bpb.de/die-bpb/presse/pressemitteilungen/549425/sinus-jugendstudie-2024-wie-ticken-jugendliche/> [Zuletzt geprüft am 01.11.2024].
- Burmester, M. & Mast, M. (2010). Repeated Web Page Visits and the Scanpath Theory: A Recurrent Pattern Detection Approach. *Journal of eye movement research*, 3(4). <https://doi.org/10.16910/jemr.3.4.5>
- Byrsell, F., Claesson, A., Ringh, M., Svensson, L., Jonsson, M., Nordberg, P., Forsberg, S., Hollenberg, J. & Nord, A. (2021). Machine learning can support dispatchers to better and faster recognize out-of-hospital cardiac arrest during emergency calls: A retrospective study. *Resuscitation*, 162, 218–226. <https://doi.org/10.1016/j.resuscitation.2021.02.041>
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, 56(2), 81–105. <https://doi.org/10.1037/h0046016>
- Crosby, M., Scholtz, J. & Wiedenbeck, S. (2002). The roles beacons play in comprehension for novice and expert programmers. In: Kuljis, J., Baldwin, L., Scoble, R. *14th Workshop of the Psychology of Programming Interest Group* (S. 58-73).
- Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A. & Sharit, J. (2006). Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging*, 21(2), 333–352. <https://doi.org/10.1037/0882-7974.21.2.333>
- Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results [Doctoral Theses]*. <https://dspace.mit.edu/handle/1721.1/15192>
- DeGEval – Gesellschaft für Evaluation. (2016). *DeGEval-Standards DeGEval-Standards für Evaluation*. https://www.degeval.org/fileadmin/content/Z03_Publikationen/DeGEval-Standards/DeGEval-Standards_fuer_Evaluation.pdf [Zuletzt geprüft am 01.11.2024].
- Dong, X., Ding, F., Zhou, S., Ma, J., Li, N., Maimaitiming, M., Xu, Y., Guo, Z., Jia, S., Li, C., Luo, S., Bian, H., Luobu, G., Yuan, Z., Shi, H., Zheng, Z.-J., Jin, Y. & Huo, Y. (2022). Optimizing an Emergency Medical Dispatch System to Improve Prehospital Diagnosis and Treatment of Acute Coronary Syndrome: Nationwide Retrospective Study in China. *Journal of medical Internet research*, 24(11), e36929. <https://doi.org/10.2196/36929>
- Döring, N., Bortz, J., Pöschl, S., Werner, C. S., Schermelleh-Engel, K., Gerhard, C. & Gäde, J. C. (2015). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Springer-Lehrbuch*. Springer.
- Dörner, D. (2008). Umgang mit Komplexität. In A. von Gleich & S. Gößling-Reisemann (Hrsg.), *Industrial Ecology* (S. 284–302). Vieweg+Teubner. https://doi.org/10.1007/978-3-8351-9225-6_24
- Duchowski, A. T. (2003). *Eye Tracking Methodology: Theory and Practice*. Springer London. <https://doi.org/10.1007/978-1-4471-3750-4>

- Egner, S., Reimann, S., Hoeger, R. & Zangemeister, W. H. (2018). Attention and Information Acquisition: Comparison of Mouse-Click with Eye-Movement Attention Tracking. *Journal of eye movement research*, 11(6). <https://doi.org/10.16910/jemr.11.6.4>
- Elsenbast, C. (2022). *Fragebogen - Technology commitment of high responsibility teams [Dataset]*. FORDATIS. <https://doi.org/10.24406/fordatis/197>
<https://doi.org/10.24406/fordatis/197>
- Elsenbast, C. (2023). *TAKAI - Beispielitems und Konzeptualisierung*. http://dx.doi.org/10.24406/fordatis/271* [Zuletzt geprüft am 01.11.2024].
- Elsenbast, C. (2024). *Attitude of Emergency Dispatchers Towards Artificial Intelligence – A Black Box of Expectations*. <https://doi.org/10.24406/fordatis/363>
- Elsenbast, C., Hofmann, T. & Böhm, D. (2023). 2. Studie zur Berufstreue angehender Notfallsanitäter:innen - Berufstreuestudie 2 (BTS2). https://www.dgre.org/wp-content/uploads/2024/08/Berufstreuestudie-2-Finale-Version.pdf* [Zuletzt geprüft am 20.09.2024].
- Emami, P. & Javanmardi, K. (2023). Enhancing Emergency Response through Artificial Intelligence in Emergency Medical Services Dispatching; a Letter to Editor. *Archives of academic emergency medicine*, 11(1), e60. <https://doi.org/10.22037/aaem.v11i1.2097>
- Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Fakhoury, S., Ma, Y., Arnaoudova, V. & Adesope, O. (2018). The effect of poor source code lexicon and readability on developers' cognitive load. In F. Khomh, C. K. Roy & J. Siegmund (Hrsg.), *Proceedings of the 26th Conference on Program Comprehension* (S. 286–296). ACM. <https://doi.org/10.1145/3196321.3196347>
- Festinger, L. (1957). *A theory of cognitive dissonance. A theory of cognitive dissonance*. Stanford University Press.
- Franklin, A., Gantela, S., Shifarraw, S., Johnson, T. R., Robinson, D. J., King, B. R., Mehta, A. M., Maddow, C. L., Hoot, N. R., Nguyen, V., Rubio, A., Zhang, J. & Okafor, N. G. (2017). Dashboard visualizations: Supporting real-time throughput decision-making. *Journal of biomedical informatics*, 71, 211–221. <https://doi.org/10.1016/j.jbi.2017.05.024>
- Fritz, T., Begel, A., Müller, S. C., Yigit-Elliott, S. & Züger, M. (2014). Using psycho-physiological measures to assess task difficulty in software development. In P. Jalote, L. Briand & A. van der Hoek (Hrsg.), *Proceedings of the 36th International Conference on Software Engineering* (S. 402–413). ACM. <https://doi.org/10.1145/2568225.2568266>
- Furangi, A., Contiero, G., Olola, C., Esposito, S., Ferlito, S., Ferrari, F., Sanna, R. & Bermano, F. (2021). Is There any Correlation between Over/Under Triage and Number of Consecutive Working Houes in the Emergency Medical Communication Center? *Annals of Emergency Dispatch & Response*, 9(1).
- Gegenfurtner, A., Eichinger, A., Latzel, R., Dietrich, M. P., Barkowsky, M., Glufke, A., Stadler, A. & Stern, W. (2018). Mobiles Eye-Tracking in den angewandten Wissenschaften. Vorab-Onlinepublikation. <https://doi.org/10.25929/bjas.v4i1.54> (Bavarian Journal of Applied Sciences, 4 (2018), 370-395).
- Glaser, B. G. & Strauss, A. L. (2017). *The Discovery of Grounded Theory*. Routledge. <https://doi.org/10.4324/9780203793206>

- Gläser, J. & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse* (4. Aufl.). VS Verlag für Sozialwissenschaften | Springer Fachmedien Wiesbaden GmbH.
- Godfroid, A. & Hui, B. (2020). Five common pitfalls in eye-tracking research. *Second Language Research*, 36(3), 277–305. <https://doi.org/10.1177/0267658320921218>
- Göransson, B., Gulliksen, J. & Boivie, I. (2003). The usability design process – integrating user-centered systems design in the software development process. *Software Process: Improvement and Practice*, 8(2), 111–131. <https://doi.org/10.1002/spip.174>
- Große, K. (2020). Vorsicht vor der Pseudo-Menschenzentrierung. *Tagesspiegel Background*(Online). <https://background.tagesspiegel.de/digitalisierung-und-ki/briefing/vorsicht-vor-der-pseudo-menschenzentrierung>
- Hagemann, V., Kluge, A. & Ritzmann, S. (2011). High Responsibility Teams – Eine systematische Analyse von Teamarbeitskontexten für einen effektiven Kompetenzerwerb. *Psychologie des Alltagshandelns*, 4, 22–42.
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology. Human Mental Workload* (Bd. 52, S. 139–183). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Heinecke-Müller, M. (2021). *Überzeugungssystem, Glaubenssystem* https://dorsch.hogrefe.com/stichwort/ueberzeugungssystem-glaubenssystem*].
- Herbig, B. & Müller, A. (2016). Hohe Belastungen in einer integrierten Rettungsleitstelle. *Neurotransmitter*, 27(9), 12–18.
- Hessels, R. S. & Hooge, I. T. C. (2019). Eye tracking in developmental cognitive neuroscience - The good, the bad and the ugly. *Developmental cognitive neuroscience*, 40, 100710. <https://doi.org/10.1016/j.dcn.2019.100710>
- Hobbs, S. D. (2002). Measuring nurses' computer competency: an analysis of published instruments. *Computers, informatics, nursing : CIN*, 20(2), 63–73. <https://doi.org/10.1097/00024665-200203000-00012>
- Hofinger, G. (2003). Entscheiden in komplexen Situationen – Anforderungen und Fehler. In R. Heimann, S. Strohschneider & H. Schaub (Hrsg.), *Schriftenreihe der Plattform Menschen in komplexen Arbeitswelten e.V. Entscheiden in kritischen Situationen: Neue Perspektiven und Erkenntnisse* (S. 3–22). Verl. für Polizeiwiss.
- Hofinger, G. (2017). Entscheidungsfindung. In A. Hackstein, V. Hagemann, F. von Kaufmann & H. Regener (Hrsg.), *Handbuch Simulation* (S. 104–106). S+K, Verlagsgesellschaft Stumpf + Kossendey mbH.
- Hofmann, T. & Macke, M. (2020). *Berufstreue von angehenden Notfallsanitäter*innen - Eine Befragung von Auszubildenden über ihren Berufsverbleib*. Deutsche Gesellschaft für Rettungswissenschaften e.V. https://www.dgre.org/download/508/* [Zuletzt geprüft am 01.11.2024].
- Holmqvist, K., Nyström, M. & Mulvey, F. (2012). Eye tracker data quality. In C. H. Morimoto, H. Istance, S. N. Spencer, J. B. Mulligan & P. Qvarfordt (Hrsg.), *Proceedings of the Symposium on Eye Tracking Research and Applications* (S. 45–52). ACM. <https://doi.org/10.1145/2168556.2168563>
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>

- Jacob, R. J. & Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research. In *The Mind's Eye* (S. 573–605). Elsevier. <https://doi.org/10.1016/B978-044451020-4/50031-1>
- Jones, J. W. (2009). Selection of Grounded Theory as an Appropriate Research Methodology for a Dissertation: One Student's Perspective. *The Grounded Theory Review: An international journal*, 8(2), 23–34.
- Karn, K. S., Ellis, S. & Juliano, C. (1999). The hunt for usability. In M. E. Atwood (Hrsg.), *CHI '99 extended abstracts on Human factors in computing systems - CHI '99* (S. 173). ACM Press. <https://doi.org/10.1145/632716.632823>
- Karnath, H.-O. & Thier Peter (Hrsg.). (2012). *Springer-Lehrbuch. Kognitive Neurowissenschaften* (3. Aufl.). Springer.
- Kassner, K. & Wassermann, P. (2002). Nicht überall, wo Methode draufsteht, ist auch Methode drin: Zur Problematik der Fundierung von ExpertInneninterviews. In A. Bogner, B. Littig & W. Menz (Hrsg.), *Das Experteninterview - Theorie, Methode, Anwendung* (S. 95–113). Springer Fachmedien Wiesbaden GmbH.
- Kerres, M. (2018). *Mediendidaktik: Konzeption und Entwicklung digitaler Lernangebote* (5. Aufl.). *De Gruyter Studium*. De Gruyter Oldenbourg.
- Keunecke, J. G., Gall, C., Birkholz, T., Moritz, A., Eiche, C. & Prottengeier, J. (2019). Workload and influencing factors in non-emergency medical transfers: a multiple linear regression analysis of a cross-sectional questionnaire study. *BMC health services research*, 19(1), 812. <https://doi.org/10.1186/s12913-019-4638-4>
- Kollmann, T., Stöckmann, C., Kensbock, J. M. & Peschl, A. (2020). What satisfies younger versus older employees, and why? An aging perspective on equity theory to explain interactive effects of employee age, monetary rewards, and task contributions on job satisfaction. *Human Resource Management*, 59(1), 101–115. <https://doi.org/10.1002/hrm.21981>
- Koo, T. K. & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Krafft, T., Neuerer, M., Böbel, S. & Reuter-Oppermann, M. (2022). *Notfallversorgung & Rettungsdienst in Deutschland. Partikularismus vs. Systemdenken*. Bertelsmann-Stiftung.
- Kramer, A. F., Weber, T. A. & Watson, S. E. (1997). Object-based attentional selection--grouped arrays or spatially invariant representations? comment on vecera and Farah (1994). *Journal of experimental psychology. General*, 126(1), 3–13. <https://doi.org/10.1037/0096-3445.126.1.3>
- Krugmann, L., Ben-Amar, M., Boden, A., Hagemann, V., Kluge, A., Ontrup, G. & Kaufmann, F. von (2021). Modernes Simulationstraining in der Integrierten Leitstellen München: Teamprozesse erfolgreich optimieren. *BOS Leitstelle aktuell*, 11, 193-197.
- Krummenacher, J. & Müller, H. (o. D.). *Aufmerksamkeit, ortsbasierte* <https://dorsch.hogrefe.com/stichwort/aufmerksamkeit-ortsbasierte>. https://dorsch.hogrefe.com/stichwort/aufmerksamkeit-ortsbasierte* [Zuletzt geprüft am 01.11.2024].
- Lee, S., Hooshyar, D., Ji, H., Nam, K. & Lim, H. (2018). Mining biometric data to predict programmer expertise and task difficulty. *Cluster Computing*, 21(1), 1097–1107. <https://doi.org/10.1007/s10586-017-0746-2>

- Lincoln, Y. S., Guba, E. G. & Pilotta, J. J. (1985). Naturalistic inquiry. *International Journal of Intercultural Relations*, 9(4), 438–439. [https://doi.org/10.1016/0147-1767\(85\)90062-8](https://doi.org/10.1016/0147-1767(85)90062-8)
- Luiz, T., Marung, H., Pollach, G. & Hackstein, A. (2019). Implementierungsgrad der strukturierten Notrufabfrage in deutschen Leitstellen und Auswirkungen ihrer Einführung [Degree of implementation of structured answering of emergency calls in German emergency dispatch centers and effects of the introduction in daily practice]. *Der Anaesthetist*, 68(5), 282–293. <https://doi.org/10.1007/s00101-019-0570-6>
- Maletzki, C., Grumbach, L., Rietzke, E. & Bergmann, R. (2023). Towards Hybrid Intelligent Support Systems for Emergency Call Handling. In A. Martin, K. Hinkelmann, H.-G. Fill, A. Gerber, D. Lenat, R. Stolle & F. van Harmelen (Hrsg.), *Proceedings of the AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023)*.
- Manzi, S., Reuter-Oppermann, M., Rachuba, S. & Morana, S. (2018). Assessing Information Requirements for Complex Decision Making in Healthcare. *26th European Conference on Information Systems 2018*
- Marung, H., Strametz, R., Roesner, H., Reifferscheid, F., Petzina, R., Klemm, V., Trifunovic-Koenig, M. & Bushuven, S. (2023). Second Victims among German Emergency Medical Services Physicians (SeViD-III-Study). *International Journal of Environmental Research and Public Health*, 20(5), 4267. <https://doi.org/10.3390/ijerph20054267>
- Mayr, B. (2020). Strukturierte bzw. standardisierte Notrufabfrage: Leisten die Systeme tatsächlich, was sie vorgeben zu leisten? [Structured or protocol-based dispatching programs]. *Notfall & rettungsmedizin*, 23(7), 505–512. <https://doi.org/10.1007/s10049-020-00733-4>
- Mays, N. & Pope, C. (2000). Qualitative research in health care. Assessing quality in qualitative research. *BMJ (Clinical research ed.)*, 320(7226), 50–52. <https://doi.org/10.1136/bmj.320.7226.50>
- McCallin, A. M. (2003). Designing a grounded theory study: some practicalities. *Nursing in critical care*, 8(5), 203–208. <https://doi.org/10.1046/j.1362-1017.2003.00033.x>
- Min-Allah, N., Jan, F. & Alrashed, S. (2021). Pupil detection schemes in human eye: a review. *Multimedia Systems*, 27(4), 753–777. <https://doi.org/10.1007/s00530-021-00806-5>
- Minh, D., Wang, H. X., Li, Y. F. & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, 55(5), 3503–3568. <https://doi.org/10.1007/s10462-021-10088-y>
- Mollenkopf, H. & Kaspar, R. (2004). Technisierte Umwelten als Handlungs- und Erlebensräume älterer Menschen. In G. M. Backes, W. Clemens & H. Künemund (Hrsg.), *Lebensformen und Lebensführung im Alter* (S. 193–221). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-10615-9_10
- Müller, H. J. & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of experimental psychology. Human perception and performance*, 15(2), 315–330. <https://doi.org/10.1037/0096-1523.15.2.315>
- Müller, H. & Krummenacher, J. (2012). Funktionen und Modelle der visuellen Aufmerksamkeit. In H.-O. Karnath & Thier Peter (Hrsg.), *Springer-Lehrbuch. Kognitive Neurowissenschaften* (3. Aufl., S. 307–321). Springer.
- Mullin, J., Anderson, A. H., Smallwood, L., Jackson, M. & Katsavras, E. (2001). Eye-Tracking Explorations in Multimedia Communications. In A. Blandford, J. Vanderdonck & P. Gray

- (Hrsg.), *People and Computers XV—Interaction without Frontiers* (S. 367–382). Springer London. https://doi.org/10.1007/978-1-4471-0353-0_22
- Nationale Armutskonferenz. (2017). *Soziale Teilhabe durch Arbeit - Sozialpolitische Anforderungen an die Beschäftigungsförderung Diskussionsgrundlage und Forderungen*. <https://www.nationale-armutskonferenz.de/wp-content/uploads/2017/08/nak-soziale-Teilhabe-durch-Arbeit-12-2014.pdf>* [Zuletzt geprüft am 01.11.2024].
- Nerdinger, F. W., Blickle, G., Schaper, N. & Solga, M. (2018). *Arbeits- und Organisationspsychologie. Springer-Lehrbuch*. Springer.
- Neyer, F. J., Felber, J. & Gebhardt, C. (2012). Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. *Diagnostica*, 58(2), 87–99. <https://doi.org/10.1026/0012-1924/a000067>
- Oldenburg, M., Wilken, D., Wegner, R., Poschadel, B. & Baur, X. (2014). Job-related stress and work ability of dispatchers in a metropolitan fire department. *Journal of Occupational Medicine and Toxicology*, 9(1). <https://doi.org/10.1186/s12995-014-0031-8>
- Oyekoya, O. & Stentiford, F. (2005). A performance comparison of eye tracking and mouse interfaces in a target image identification task. In *2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT 2005)* (S. 139–144). IET. <https://doi.org/10.1049/ic.2005.0723>
- Parasuraman, R., Sheridan, T. B. & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE transactions on systems, man, and cybernetics. Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society*, 30(3), 286–297. <https://doi.org/10.1109/3468.844354>
- Pentzold, C., Bischof, A. & Heise, N. (2018). *Praxis Grounded Theory*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-15999-3>
- Pergert, P. (2009). Methodological Learning-by-doing: Challenges, lessons learned and rewards. *The Grounded Theory Review: An international journal*, 8(2), 65–75.
- Perthes, A. & Jakob, T. (2012). Weisungsrecht und Ladungssicherung im Rettungsdienst. *Der Notarzt*, 28(02), 57–62. <https://doi.org/10.1055/s-0031-1298972>
- Phillips-Wren, G. & Adya, M. (2020). Decision making under stress: the role of information overload, time pressure, complexity, and uncertainty. *Journal of Decision Systems*, 29(sup1), 213–225. <https://doi.org/10.1080/12460125.2020.1768680>
- Posner, M. I., Snyder, C. R. & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology*, 109(2), 160–174.
- Prottengeier, J., Keunecke, J. G., Gall, C., Eiche, C., Moritz, A. & Birkholz, T. (2019). Single mission workload and influencing factors in German prehospital emergency medicine - a nationwide prospective survey of 1361 emergency missions. *Scandinavian journal of trauma, resuscitation and emergency medicine*, 27(1), 75. <https://doi.org/10.1186/s13049-019-0650-2>
- Pruitt, C. M. & Liebelt, E. L. (2010). Enhancing patient safety in the pediatric emergency department: teams, communication, and lessons from crew resource management. *Pediatric emergency care*, 26(12), 942-8; quiz 949-51. <https://doi.org/10.1097/PEC.0b013e3181fec9cf>
- Pupil Labs GmbH. (o. D.). *Core*. <https://pupil-labs.com/products/core>* [Zuletzt geprüft am 01.11.2024].

- Raehlmann, I. (2017). Voraussetzungen der Entwicklung und Anwendung von Technik im Arbeitsprozess. *Zeitschrift für Arbeitswissenschaft*, 71(2), 120–127. <https://doi.org/10.1007/s41449-017-0053-z>
- Raschke, M. (2024). *Überprüfung der Datenqualität (Teil 1)* [Video]. Youtube. <https://www.youtube.com/watch?v=1sRjYLSn2F4>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Reader, T. W., Flin, R., Mearns, K. & Cuthbertson, B. H. (2007). Interdisciplinary communication in the intensive care unit. *British journal of anaesthesia*, 98(3), 347–352. <https://doi.org/10.1093/bja/ael372>
- Reuter-Oppermann, M. & Richards, D. (2019). Decision Support for EMS Policy Making Using Data Analytics and Real-Time Alerts. In *2019 IEEE World Congress on Services (SERVICES)* (S. 266–271). IEEE. <https://doi.org/10.1109/SERVICES.2019.00079>
- Rivollier, G., Quinton, J.-C., Gonthier, C. & Smeding, A. (2021). Looking with the (computer) mouse: How to unveil problem-solving strategies in matrix reasoning without eye-tracking. *Behavior research methods*, 53(3), 1081–1096. <https://doi.org/10.3758/s13428-020-01484-3>
- Santini, T., Fuhl, W. & Kasneci, E. (2017). CalibMe. In G. Mark, S. Fussell, C. Lampe, m. schraefel, J. P. Hourcade, C. Appert & D. Wigdor (Hrsg.), *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (S. 2594–2605). ACM. <https://doi.org/10.1145/3025453.3025950>
- Schepers, J. & Wetzels, M. (2007). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management*, 44(1), 90–103. <https://doi.org/10.1016/j.im.2006.10.007>
- Schnipke, S. K. & Todd, M. W. (2000). Trials and tribulations of using an eye-tracking system. In M. Tremaine (Hrsg.), *CHI '00 Extended Abstracts on Human Factors in Computing Systems* (S. 273–274). ACM. <https://doi.org/10.1145/633292.633452>
- Schnotz, W. & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, 19(4), 469–508. <https://doi.org/10.1007/s10648-007-9053-4>
- Scholz, M. L., Collatz-Christensen, H., Blomberg, S. N. F., Boebel, S., Verhoeven, J. & Krafft, T. (2022). Artificial intelligence in Emergency Medical Services dispatching: assessing the potential impact of an automatic speech recognition software on stroke detection taking the Capital Region of Denmark as case in point. *Scandinavian journal of trauma, resuscitation and emergency medicine*, 30(1), 36. <https://doi.org/10.1186/s13049-022-01020-6>
- Sharafi, Z., Sharif, B., Guéhéneuc, Y.-G., Begel, A., Bednarik, R. & Crosby, M. (2020). A practical guide on conducting eye tracking studies in software engineering. *Empirical Software Engineering*, 25(5), 3128–3174. <https://doi.org/10.1007/s10664-020-09829-4>
- Sharafi, Z., Soh, Z. & Guéhéneuc, Y.-G. (2015). A systematic literature review on the usage of eye-tracking in software engineering. *Information and Software Technology*, 67, 79–107. <https://doi.org/10.1016/j.infsof.2015.06.008>
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>

- Starcke, K. & Brand, M. (2012). Decision making under stress: a selective review. *Neuroscience and biobehavioral reviews*, 36(4), 1228–1248. <https://doi.org/10.1016/j.neubiorev.2012.02.003>
- Steiner, E. & Benesch, M. (2018). *Der Fragebogen: Von der Forschungsidee zur SPSS-Auswertung* (5. aktual. u. überarb. Auflage). UTB.
- Steinke, I. (2019). Güterkriterien qualitativer Forschung. In U. Flick, E. von Kardorff & I. Steinke (Hrsg.), *Qualitative Forschung - Ein Handbuch (translated: Qualitative Research - A Handbook)* (S. 319–331). Rohwoldt Taschenbuch Verlag.
- Stieglitz, R.-D. (2014). Psychometrische Verfahren in der Psychotherapie. *Verhaltenstherapie*, 24(1), 56–65. <https://doi.org/10.1159/000358912>
- Strutz, N., Kuntz, S., Lahmann, N. & Steinert, A. (2020). Analyse der Technikbereitschaft und -nutzung von Pflegeinnovationstechnologien von Mitarbeiter*innen im Pflegeprozess. *HeilberufeScience*, 11(3-4), 27–34. <https://doi.org/10.1007/s16024-020-00339-3>
- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2011). Cognitive Load Theory. In J. Mestre & B. H. Ross (Hrsg.), *The Psychology of Learning and Motivation: Cognition in Education* (S. 37–77).
- Tobii. (2023). *Dark and Bright pupil tracking*. https://connect.tobii.com/s/article/What-is-dark-and-bright-pupil-tracking?language=en_US* [Zuletzt geprüft am 01.11.2024].
- Trautmann, R. & Ballé, J. (2022). *EXPECT - Eine Studie über Erwartungen, Einstellungen und Erfahrungen mit Leitstellen der nichtpolizeilichen Gefahrenabwehr in der Bundesrepublik Deutschland*. https://www.dgre.org/wp-content/uploads/2022/09/Studie-Expect-mit-Umschlag.pdf* [Zuletzt geprüft am 01.11.2024].
- Trautmann, R., Reuter-Oppermann, M. & Christiansen, J. (2022). *PSAP-G-ONE - Eine explorativ-deskriptive Studie über Leitstellen der nichtpolizeilichen Gefahrenabwehr in der Bundesrepublik Deutschland*. https://www.dgre.org/download/2067/* [Zuletzt geprüft am 01.11.2024].
- Treue, S., Womelsdorf, T. & Martinez-Trujillo, J. C. (2007). Visuelle Aufmerksamkeit: Von Orten, Eigenschaften und Kontrasten. *e-Neuroforum*, 13(2), 55–61. <https://doi.org/10.1515/nf-2007-0204>
- Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological bulletin*, 63, 384–399.
- Usó-Doménech, J. L. & Nescolarde-Selva, J. (2016). What are Belief Systems? *Foundations of Science*, 21(1), 147–152. <https://doi.org/10.1007/s10699-015-9409-z>
- van Renswoude, D. R., Raijmakers, M. E. J., Koornneef, A., Johnson, S. P., Hunnius, S. & Visser, I. (2018). Gazepath: An eye-tracking analysis tool that accounts for individual differences and data quality. *Behavior research methods*, 50(2), 834–852. <https://doi.org/10.3758/s13428-017-0909-3>
- Venkatesh, V. & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186–204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Westerheide, J. E. & Schott, A. (2022). Kritische Wissenschaft im Elfenbeinturm oder Gestaltungswissenschaft im Betrieb? *Zeitschrift für Soziologie*, 51(4), 335–349. <https://doi.org/10.1515/zfsoz-2022-0024>

- Wiczorek, R. & Meyer, J. (2019). Effects of Trust, Self-Confidence, and Feedback on the Use of Decision Automation. *Frontiers in psychology*, 10, 519. <https://doi.org/10.3389/fpsyg.2019.00519>
- Xie, S. L. (2009). Striking a Balance between Program Requirements and GT Principles: Writing a compromised GT proposal. *The Grounded Theory Review: An international journal*, 8(2), 35–47.
- Yadav, D. (2022). Criteria for Good Qualitative Research: A Comprehensive Review. *The Asia-Pacific Education Researcher*, 31(6), 679–689. <https://doi.org/10.1007/s40299-021-00619-0>
- Yigitbasioglu, O. M. & Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1), 41–59. <https://doi.org/10.1016/j.accinf.2011.08.002>

Anlage 1 - Eigenständigkeitserklärung

Hiermit erkläre ich, Christian Elsenbast, dass ich die Arbeit ohne unerlaubte Hilfe angefertigt habe, keine anderen, als die angegebenen Quellen und Hilfsmittel benutzt habe und dass ich die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Einer Überprüfung der Dissertation mit qualifizierter Software im Rahmen der Untersuchung von Plagiatsvorwürfen stimme ich zu.

Christian Elsenbast, Hochspeyer den 03.11.2024

Anlage 2 – Eigenleistungen

Die nachfolgende Tabelle stellt die Eigenleistungen zu Publikation 1, Publikation 4 und Publikation 5 dar. Diese wurden gemeinsam mit anderen Autor:innen verfasst. Publikation 2, Publikation 3 und Publikation 6 wurden in alleiniger Autorschaft angefertigt.

Tabelle 4: Beschreibung der Eigenleistungen

Lfd.-Nr.	Publikation	Anteil der Eigenleistung	Eigenanteil
P1	Elsenbast, C., & Hagemann, V. (2023); <i>Technology commitment of emergency medical service practitioners and dispatchers</i> , International Paramedic Practice, Vol 13, No 3, DOI: 10.12968/ippr.2023.13.3.59	75%	Psychometrische Weiterentwicklung der Kurzskala Technikbereitschaft, Erstellung, Testung und Rollout des Fragebogens, deskriptiv- und inferenzstatische Auswertung mit Validierung der weiterentwickelten Skala durch konfirmatorische Faktorenanalyse, Manuskripterstellung und Einreichung.
P4	Elsenbast, C., Reuter-Oppermann, M. & Dahlmann, P.; <i>Attitude of Emergency Dispatchers Towards Artificial Intelligence - A Black Box of Expectations [under review]</i>	75%	Iterative Weiterentwicklung des semistrukturierten Interview-Leitfadens, Führen der Interviews, Transkription, Anwendung der Grounded Theory, Theoriebildung, Organisation und Durchführung der Interviewvalidierung, Organisation der gemeinsamen Manuskripterstellung, aktive Teilnahme am Schreibprozess, Einreichung.
P5	Maletzki, C., Elsenbast, C. & Reuter-Oppermann, M. (2023); <i>Towards Human-AI Interaction in Medical Emergency Call Handling</i> , 57. Hawaii International Conference on System Sciences 2024, URI: https://hdl.handle.net/10125/106790	60%	Verantwortliche inhaltliche Gestaltung des Kapitels „Measuring Mental Workload in Emergency Call Handling“ mit der dazugehörigen Beschreibung der Datenlagen aus den Bezugswissenschaften, Anfertigung von repräsentativen Bildern und Ablaufschemata der Feldstudie, Mitwirkung bei der Erstellung von „Introduction“ sowie „Conclusion and Outlook“, sowie aktive Teilnahme an Diskussionen zu den weiteren Kapiteln.

Anlage 3 – Elaboration Expert:innen-Interview

In diesem Abschnitt wird der Begriff „Expert:innen“ als methodologische Konkretisierung zur Beantwortung von RQ 2 vertiefend erörtert.

Die Anwendung von Expert:innen-Interviews in der empirischen Sozialforschung wurde bereits kritisch betrachtet (Kassner & Wassermann, 2002, S. 95). Kassner und Wassermann heben hervor, dass es ebenso wenig ein „Hausfraueninterview“ oder „Beamteninterview“ gebe (Bogner et al., 2002, S. 39). Sie kritisieren weiterhin, dass trotz der erfolgreichen Nutzung von Expert:innen-Interviews eine mangelnde Reflexion zu beklagen sei. Sie argumentieren, dass Expert:innen-Interviews nur dann als eine eigenständige Methode gelten können, wenn sie als ein über das einzelne Forschungsvorhaben hinaus anwendbares Werkzeug verstanden werden. Die Generalisierbarkeit dieser Methode stellen Kassner und Wassermann (2002, S. 107–108) in Frage und führen an, dass sich keine spezifischen Vorschläge zur Gestaltung von Interviews finden ließen, die nicht generell auf qualitative Interviews anwendbar wären, was eine deutliche Abgrenzung vermissen lässt.

Daraus ergibt sich, dass Kassner und Wassermann auf die Kontextsensitivität von Expert:innen-Interviews aufmerksam machen. Diese Spezifität konstituiert eine zusätzliche Differenzierung gegenüber narrativen und problemzentrierten Interviews, welche aufgrund ihrer ausgeprägten metatheoretischen Fundierung flexibel und ohne signifikante Anpassungsleistungen in diversen Forschungskontexten und bei unterschiedlichen Zielgruppen eingesetzt werden können.

Demzufolge leitet sich die Besonderheit des Expert:innen-Interviews nicht aus methodenspezifischen Verfahren ab, sondern resultiert aus der essentiellen, forschungskontextbezogenen Integration in das spezifische Forschungsdesign. Diese Kontextabhängigkeit bedingt für Forschende die Notwendigkeit, den Terminus des Experten bzw. der Expertin in jedem Forschungsvorhaben individuell zu redefinieren und in Einklang mit dem Forschungsdesign zu bringen.

Bezüglich des Verständnisses des Expert:innenbegriffs folgt diese Arbeit der Auffassung von Gläser und Laudel (2010), welche die soziale Dimension des Expert:innenbegriffs betonen, indem sie Expert:innen als Personen definieren, die über spezifisches Wissen in sozialen Belangen verfügen und somit „Expertenwissen“ von „Alltagswissen“ und dem Wissen spezialisierter Laien differenzieren. Bogner et al. (2014, S. 13–14) argumentieren, dass das spezifische Spezialwissen von Expert:innen innerhalb ihres professionellen Kontextes ihnen

Macht verleiht, was ein motivierender Faktor für die Durchführung von Expert:innen-Interviews ist. Dies impliziert die Notwendigkeit einer Selbstreflexion der Forschenden über ihre Subjektivität und Perspektive. Expert:innen werden von Bogner et al. (2014, S.13) definiert als Personen, die basierend auf spezifischem Praxis- oder Erfahrungswissen, das sich auf einen klar definierten Problembereich bezieht, die Möglichkeit geschaffen haben, ihr Handlungsfeld sinnhaft und handlungsleitend für andere zu strukturieren. Die Abgrenzung von Expert:innen zu Eliten erfolgt über unterschiedliche Kriterien: Während bei Expert:innen kognitive Ressourcen im Vordergrund stehen, sind es bei Eliten Kontakte, Netzwerke und ein spezifischer Habitus. Expert:innen können jedoch Teil einer Funktionselite sein. Die Unterscheidung zu Spezialist:innen, die zwar über spezifisches Fachwissen verfügen, aber nicht die Fähigkeit haben, ihre Interpretationen handlungsleitend einzubringen, wird ebenfalls hervorgehoben.

Wird die Verortung von Expert:innen- und Spezialist:innen-Interviews elaboriert, kann festgehalten werden, dass die charakteristischen Merkmale des Expert:innen-Interviews nicht durch eine spezifische methodologische Praxis determiniert werden, noch beansprucht das Expert:innen-Interview eine explizite methodische Eigenständigkeit (Kassner & Wassermann, 2002, S. 108), wie dies beispielsweise bei dem problemzentrierten Interview der Fall ist (Witzel & Reiter, 2012 nach Bogner et al., 2014, S. 9), dem episodischen Interview (Flick, 2011 nach Bogner et al., 2014, S. 9) oder auch dem narrativen Interview (Schütze, 1983 nach Bogner et al., 2014, S. 9). Demnach lässt sich das Expert:innen-Interview eher als ein spezifisches Forschungsinteresse verstehen und weniger als eine standardisierte methodische Herangehensweise an sich (Kassner & Wassermann, 2002, S. 106). Dieses Forschungsinteresse zeichnet sich etwa durch einen erheblichen Bezug zu den Orientierungen und Interpretationsschemata einer als in ihren Handlungskapazitäten privilegierten Personengruppe aus (Kassner & Wassermann, 2002, S. 106) und identifiziert somit die Expert:innen als das Kernstück, wie auch Bogner et al. (2014, S. 9) in einem anderen Zusammenhang ausführen.

Bei den im Rahmen dieser Arbeit durchgeführten Interviews von Leitstellendisponent:innen ohne Führungsaufgaben und weiteren Positionen oder sonstigen soziale Rahmungen, handelt es sich also im engeren Sinne um Spezialist:innen. Hier interviewte Führungskräfte, wie Leiter von Leitstellen und Behörden, sind den Expert:innen zuzuordnen. Der Arbeitsmodus dieses Teilbereichs lautet in dieser Arbeit daher: Interviews von Expert:innen und Spezialist:innen.

Anlage 4 – Ergebnistabelle TAKAI (P2 und P3)

Die dargestellte Tabelle dient der detaillierten Ergebnisdarstellung zu P2 und P3.

Tabelle 5: Ergebnisse der Itemgruppen und Intraklassen-Korrelation

MW,SD, Itemgruppen, ICC	MW SD	Notfall- sanitäter: innen	Leit- stellen- Personal	Notärzt: innen	Berufs- feuerwehr- Angehörige	Pflege- fachkräfte	ICC
Intransparenz (K)	MW	3,30 ²	3,38	3,71 ^{1,2}	3,51	3,23 ¹	0,93
	SD	0,70	0,73	0,70	0,67	0,50	
Vernetztheit: Abteilungen (K)	MW	3,89	3,64 ¹	3,782	4,21 ^{1,2}	3,93	0,76
	SD	0,72	1,10	0,93	1,03	0,76	
Vernetztheit: Informations- fluss (K)	MW	3,89	3,64 ¹	3,78	4,32 ¹	3,95	0,75
	SD	0,72	1,07	0,93	1,00	0,71	
Eigendynamik (K)	MW	4,09 ⁶	4,17 ^{3,4,5}	3,71 ^{1,2,3}	4,91 ^{1,4,5,6}	2,91 ²	0,78
	SD	0,81	0,85	1,07	1,05	0,52	
Polytelie (K)	MW	4,16 ⁴	4,26	4,22 ^{1,3}	4,21 ²	3,12 ^{1,2,3,4}	0,85
	SD	0,78	0,77	1,05	1,02	0,65	
Verzögerte Rückmeldung (K)	MW	3,47	3,14 ¹	3,48	3,67 ¹	3,14	0,91
	SD	0,61	0,97	0,65	0,91	1,26	
Hierarchie (KK)	MW	3,14	3,11	2,72	3,13	3,70	0,93
	SD	0,67	1,19	0,49	0,72	0,99	
Hierarchie: Followership (KK)	MW	4,01 ^{1,2,3}	3,40 ³	2,76 ¹	2,74 ²	3,37	0,72
	SD	1,12	1,33	0,60	1,56	1,11	
Hierarchie: Leadership (KK)	MW	4,47 ^{1,3,5}	4,57 ^{2,4}	2,83 ¹	2,65 ^{2,3}	2,21 ^{4,5}	0,91
	SD	0,74	1,23	0,80	1,55	1,32	
Geschwindigkeit der Systemveränder- ungen (KK)	MW	2,15 ^{3,5}	1,83 ^{2,4,6}	3,91 ^{1,2,3}	3,25 ^{4,5}	2,94 ^{1,6}	0,86
	SD	0,57	1,76	0,65	1,28	1,08	
Persönliche Bedrohung (KK)	MW	3,10 ^{2,4}	1,16 ⁴	3,65 ^{1,2}	3,69 ^{1,3}	3,45 ³	0,77
	SD	0,82	1,08	0,71	1,23	0,73	
Beeinträchtigung der Kommu- nikation (KK)	MW	1,57	1,67	2,13	1,95	1,59	0,81
	SD	0,76	1,83	0,98	0,77	0,67	
Umweltfaktoren (KK)	MW	2,94 ^{4,6,7}	2,07 ^{2,5,7}	0,56 ^{1,2}	3,60 ^{1,3,4}	0,53 ^{3,5,6}	0,91
	SD	0,67	1,00	0,85	0,54	0,88	
Bekanntheit der Arbeits-	MW	1,59 ^{2,4}	3,87 ^{1,4}	3,31 ^{1,2}	1,46 ³	3,20 ³	0,95

umgebung (KK)	SD	0,53	0,47	0,79	0,55	0,72	
Geschwindigkeit der Bewegung des Teams (KK)	MW	4,25 ³	4,00	4,56 ²	4,27 ¹	3,33 ^{1,2,3}	0,87
	SD	0,89	1,41	0,62	0,67	1,12	
Shared Task Mental Model (S)	MW	4,93 ³	5,19	4,81 ²	5,44 ¹	4,21 ^{1,2,3}	0,79
	SD	0,80	0,93	1,17	0,50	1,40	
Shared Team Mental Model (S)	MW	5,19	5,31	4,38	5,35 ¹	4,66 ¹	0,59
	SD	0,48	0,47	1,09	0,77	0,75	
Informationssammlung (A)	MW	5,56 ³	5,59 ^{2,4}	4,15 ^{1,2,3}	5,03 ⁴	4,49 ¹	0,93
	SD	0,76	0,65	1,20	0,86	0,87	
Aufgabenpriorisierung (A)	MW	5,37 ²	5,34 ^{1,4}	4,49 ^{1,2}	5,24 ³	4,27 ^{3,4}	0,81
	SD	0,85	0,90	1,23	0,81	1,59	
Aufgabenverteilung (A)	MW	5,37 ²	5,22 ⁴	4,12 ^{1,2}	5,13 ³	2,15 ^{1,3,4}	0,79
	SD	0,89	1,25	1,55	0,89	2,47	

Abkürzungen: (K) = Konstrukt „Komplexität“, (KK) = Konstrukt „Kontextkriterien“, (A) = Konstrukt „Adaptionserfordernisse“, (S) = Konstrukt „Shared Mental Model“

Anlage 5 – Transfer der STOPV-Maßnahmen

Die dargestellte Tabelle zeigt Praxisbeispiele für einen Transfer der STOPV-Maßnahmen auf die Leitstellen, wie sie im Diskussionskapitel bereits grundsätzlich angeführt wurden.

Tabelle 6: STOPV-Maßnahmen

	Maßnahmenhierarchie	Beispiele für die Leitstelle
S	Substitution: Komplexitätsquelle vermeiden, beseitigen, reduzieren oder Eigenschaften der Quelle verändern	Klare Dispositions- und Zuweisungskriterien durch bedarfsorientierte Patientensteuerung statt Sektorendenken
T	Technische Maßnahmen: Ergonomie des Arbeitsplatzes, Softwareergonomie	Einsetzen ergonomischer Softwareprodukte, intelligente Entscheidungsunterstützungssysteme, angemessenes Signalisieren von Anrufen, Nachrichten und Tasks nach automatischer Prioritätszuweisung unter Berücksichtigung der jeweiligen Salienz des Reizes
O	Organisatorische Maßnahmen: Strukturelle Änderungen zur Bewältigung von Komplexität	Etablieren verbindlicher Standards, wie die konsequente Anwendung der standardisierten Notrufannahme, ermöglichen angemessener Ruhepause (ggf. adaptiert an das jeweilige Anrufaufkommen), engmaschige Bildschirm-pausen
P	Persönliche Maßnahmen: Hier geht es um den Schutz der Mitarbeiter vor den Auswirkungen von Komplexität	Stressmanagement-Tools, ergonomische Arbeitsplatzgestaltung, Messung und ggf. Begrenzung von akustischen Belastungen
V	Verhaltensbezogene Maßnahmen: Schulungen und Trainings zum besseren Umgang mit Komplexität,	Zielgerichtete Qualifizierungsmaßnahmen zu komplexitätsbringenden Faktoren, konsequente und überprüfbare Anwendung der etablierten Standards

Anlage 6 – Explainable AI – Umsetzungsimpulse

Die nachfolgende Tabelle gibt als Konkretisierung zu den Einlassungen im Diskussionskapitel Umsetzungsimpulse hinsichtlich erklärbarer KI (explainable AI) in Leitstellen.

Tabelle 7: Explainable AI - Umsetzungsimpulse

	Erklärung	Beispiel im Leitstellenkontext
Erklärbarkeit vor der Modellierung	Die Erklärbarkeit vor der Modellierung bezieht sich auf eine Reihe von Datenverarbeitungsansätzen, die eingesetzt werden, um einen Einblick in die Datensätze zu erhalten.	Visualisierung mit Heatmaps zur Darstellung von Einsatzschwerpunkten in der Region oder Berechnung von Entfernungen und geschätzten Anfahrtszeiten basierend auf aktuellen Verkehrsdaten.
Interpretierbare Modelle	Ein Modell gilt als interpretierbar, wenn es von Menschen allein durch Betrachtung der Modellzusammenfassung oder der Modellparameter interpretiert werden kann.	Implementierung eines Entscheidungsbaums für die Prioritätszuweisung und regelbasierte Systeme mit klaren Regeln für die Zuweisung von Ressourcen. Visualisierung des Modell-Fundierung und Wissensbasis, z. B. als Knowledge-Graph.
Erklärbarkeit nach der Modellierung	Verbessert die Erklärbarkeit der bestehenden Modelle durch Visualisierung, textuelle Begründung, Vereinfachung und Techniken zur Relevanz von Merkmalen.	Visualisierung der Zuverlässigkeit der Entscheidung des DSS durch eine niederschwellig verstehbare Grafik („Uncertainty Graph“).

Anlage 7 - Ethische Exkursion

Während in dieser Arbeit hauptsächlich um die Endanwender:innen und ihre Arbeitsrealität fokussiert wird, werden ethische Herausforderungen höchstens als Randnotiz berücksichtigt. Dieser Abschnitt soll diese Lücke als grundlegender Impuls schließen, wenngleich er nicht den Anspruch einer abschließenden ethischen Darstellung erhebt. Vorwegnehmend sei gesagt, dass die Disponent:innen in dieser Arbeit als Bestandteil der partizipativen Entwicklung und Implementierung verstanden werden.

In der modernen Softwareentwicklung gewinnt die Berücksichtigung ethischer Aspekte zunehmend an Bedeutung. Softwareprodukte, als solches ist auch ein intelligentes Entscheidungsunterstützungstool im Leitstellenkontext zu verstehen, tragen oft implizite Werte in sich, die das Verhalten der Anwender:innen beeinflussen können, weshalb es entscheidend ist, dass Entwickler ein ausgeprägtes "techno-ethisches Urteilsvermögen" entwickeln (Zuber et al., 2022, S. 2). Ein solcher impliziter Wert kann sich im Leitstellensetting beispielsweise im klassischen Weichensteller-Problem (Welzel, 1951) verwirklichen, wenn es darum geht, welche Person bei einer Verknappung von Rettungsmitteln zuerst „bedient“ wird. Das angesprochene techno-ethische Urteilsvermögen soll dazu führen, dass technische Möglichkeiten mit ethischen Implikationen verknüpft und abgewogen werden. Um dies zu fördern, empfehlen Zuber et al. (2022, S. 3) u. a. einen teambasierter Ansatz für ethische Diskussionen. Besonders agile Entwicklungsmethoden bieten durch ihren iterativen Charakter gute Voraussetzungen, um ethische Überlegungen von Beginn an und kontinuierlich in den Entwicklungsprozess zu integrieren, wie auch Zuber et al. (2022) betonen. Dies ermöglicht einen praxisnahen Umgang mit ethischen Herausforderungen und fördert die Entwicklung einer professionellen Ethik in der Softwareentwicklung, ähnlich den ethischen Standards in anderen Berufsfeldern. Folgt man diesen Einlassungen, ist das für die zukünftige grundständige und spezifische Qualifikation der Disponent:innen von Bedeutung, denn Grundzüge der Technoethik und der agilen Methoden sind dann curricular zu verankern.

Darüber hinaus stellt sich auf einem höheren Abstraktionsniveau auch grundsätzlich die Frage nach der Notwendigkeit einer völlig neue Bereichsethik für KI, die über die Annahmen der Maschinenethik, die zu einem großen Teil die vorletzte industrielle Revolution aufgreift, hinausgeht. So argumentieren auch Joisten et al. (2022), dass bestehende technoethische Konzepte nicht ausreichen, um die komplexen ethischen Herausforderungen der heutigen Technologien zu bewältigen. Die entwickelte Theorie von Joisten et al. bietet durch die

Anwendung phänomenologischer Methoden eine Lösung, um ethische „Blindstellen“ in der Entwicklung und Anwendung digitaler Technologien frühzeitig zu erkennen. Mit ihrem Konzept zielen sie darauf ab, die Entwicklung von KI-Systemen ethisch nachhaltig zu gestalten und ethische Fehler zu vermeiden, bevor sie in realen Anwendungen zu Problemen führen. Der Anwendung dieser technoethischen Grundprinzipien ist auch hinsichtlich der KI-Entwicklungen und Implementierungen im Leitstellenkontext Rechnung zu tragen, was zum Schluss führt, dass die Einbindung von Expert:innen mit ausgeprägten technoethischen Kompetenzen angezeigt ist, beispielsweise mit der konkreten Aufgabenzuweisung einer partizipativen Technikfolgenabschätzung, die den kritischen ethischen Diskurs aufnimmt.

Anlage 8 – Tabellarische Zusammenfassung der Publikationen

Um einen am Ende dieses Kapitels einen konzisen Überblick über die Publikationen zu ermöglichen, fasst die nachfolgende Tabelle die Publikationen dieser Dissertation zusammen.

Tabelle 8: Tabellarische Zusammenfassung der Publikationen

	P1	P2	P3	P4	P5	P5
	RQ 1		RQ 2		RQ 3	
Titel	Technology commitment of emergency medical service practitioners and dispatchers	Komplexität und Adaptionserfordernisse an der Nahtstelle Leitstelle-Rettungsdienst-Feuerwehr: Eine validierte Fragebogenanalyse	Zwischen Hierarchie und geteilten mentalen Modellen: Analyse mentaler Modelle an der Nahtstelle Leitstelle-Rettungsdienst-Feuerwehr	Attitude of Emergency Dispatchers Towards Artificial Intelligence - A Black Box of Expectations	Towards Human-AI Interaction in Medical Emergency Call Handling	Eye-tracking the Lifeline: Emergency Dispatchers Visual Attention in Time-Sensitive Calls
Autor:innen	Elsenbast, C. & Hagemann, V.	Elsenbast, C.	Elsenbast, C.	Elsenbast C., Reuter-Oppermann, M. & Dahmann, P.	Maletzki, C., Elsenbast C. & Reuter-Oppermann, M.	Elsenbast, C.
Forschungsfrage(n) bzw. Ziele der Studie	<ol style="list-style-type: none"> 1. In welchem Ausmaß ist Technikbereitschaft bei Rettungsdienstpersonal, Leitstellendisponent:innen und Notärzt:innen festzustellen? 2. Können signifikante Unterschiede zwischen den Berufsgruppen festgestellt werden? 3. Können signifikante altersbezogene 	Welche Handlungserfordernisse ergeben sich aus der Wahrnehmung von Komplexitätsfaktoren sowie aus Adaptionserfordernissen der Teams an den Nahtstellen der Leitstellen?	Welche Handlungserfordernisse ergeben sich aus der Analyse der berufsbezogenen Tätigkeiten von HRT an den Nahtstellen der Integrierten Leitstellen bezüglich der Wahrnehmung von geteilten mentalen Modellen und Hierarchien?	Welche Überzeugungssysteme existieren bei in Deutschland tätigen Leitstellendisponent:innen gegenüber der Implementierung von KI-basierten Entscheidungsunterstützungssystemen in integrierten Leitstellen?	Ziel der Studie ist, ein Design für hybrid-intelligente Notrufbearbeitung zu entwickeln. Außerdem werden die notwendigen Artefakte für die Mensch-KI-Interaktion definiert und grundlegende Bestandteile eines Evaluationskonzepts beschrieben.	Welche grafischen Elemente der Benutzeroberfläche des ELS werden von Leitstellendisponent:innen während des Prozesses von der Notrufannahme bis zum Einsatzende besonders beachtet?

	Effekte festgestellt werden?				
Methode	Adressierung des Fragebogens Technikbereitschaft nach Neyer et al. (2012) und Weiterentwicklung hinsichtlich KI-Technikbereitschaft mit nachfolgender Validierung durch CFA	Adressierung des Fragebogens Team-Arbeit-Kontext nach Hagemann et. al (2011) mit nachfolgender Validierung durch Multitrait-Multimethod-Methode	Semi-strukturierte Interviews mit Expert:innen, Spezialist:innen und Fokusgruppeninterviews, Auswertung durch Grounded Theory	Transfer der Leitlinien für Mensch-KI-Interaktion nach Amershi et al. (2019) auf die Leitstellen sowie Anwendung von Evaluationsleitlinien für DSS in der Leitstelle	Mobiles Eyetracking im Leitstellenbetrieb mit Auswertung von Areas of Interest
Stichprobe	<p>N = 510</p> <p>Geschlechterverteilung: 354 (69,4 %) männlich; 125 (24,5 %) weiblich; 4 (0,8 %) nicht-binär</p> <p>Durchschnittsalter: 38,9 Jahre</p> <p>Beruflicher Hintergrund: 199 (39 %) Leitstellendisponent:innen; 184 (36,1 %) Notfallsanitäter:innen; 68 (13,3 %) Notärzt:innen</p>	<p>N = 490</p> <p>Geschlechterverteilung: 361 (74,4 %) männlich; 116 (23,9 %) weiblich; 8 (1,7 %) nicht-binär</p> <p>Durchschnittsalter: 40,6 Jahre</p> <p>Beruflicher Hintergrund: 135 (37 %) Leitstellendisponent:innen; 121 (24,7 %) Notfallsanitäter:innen; 43 (8,8 %) Pflegefachkräfte; 48 (9,8 %) Notärzt:innen; 39 (8,0 %) Berufsfeuerwehr-Beamte:innen</p>	<p>N = 31</p> <p>Geschlechterverteilung: 24 männlich; 7 weiblich; 0 nicht-binär.</p> <p>Durchschnittsalter: 39,3 Jahre</p> <p>Beruflicher Hintergrund (Mehrfachnennung): 19 Leitstellendisponent:innen; 19 Notfallsanitäter:innen; 4 Rettungssanitäter:innen; 5 Berufsfeuerwehrbeamte:innen; 2 Notärzt:innen; 6 Führungskräfte;</p>	Literaturbasiert	<p>N = 33</p> <p>Geschlechterverteilung: 23 männlich; 3 weiblich; 4 wüshten keine Geschlechtsbezeichnung.</p> <p>Durchschnittsalter: 40,5 Jahre</p> <p>Beruflicher Hintergrund: 33 Leitstellendisponent:innen davon: 24 Rettungsdiensthintergrund, 9 Feuerwehrhintergrund</p>

				2 Philosoph:innen; 1 Psycholog:in; 2 Ingenieurs- wissenschaftler:innen		
Ergebnisse	<p>Negative Korrelation zwischen Alter und Technikbereitschaft ($r=-0,32$); gute allgemeine Technikbereitschaft über alle Gruppen sowie moderate Technikbereitschaft gegenüber KI über alle Gruppen; jedoch auch statistische Ausreißer nach unten; Alter erklärt 12,48% der Varianz bei Technikbereitschaft; starke signifikante Korrelation ($r=0,70$) zwischen allgemeiner und KI-spezifischer Technikbereitschaft.</p> <p>Validierung: Gute Reliabilität und Validität.</p>	<p>Komplexitätsfaktoren: Fast alle Berufsgruppen bewerten über Trennwert; höchste Bewertung bei Vernetztheit und Eigendynamik durch Feuerwehr-Beamten:innen;</p> <p>Adaptionserfordernisse: Überschreitung des Trennwerts der Disponent:innen bei Informationsammlung, Aufgabenverteilung und Aufgabepriorisierung.</p> <p>Validierung: Gute bis exzellente Reliabilität und Validität.</p>	<p>Mentale Modelle: Überschreiten der Trennwerte von Leitstellendisponent:innen und Berufsfeuerwehr-Beamten:innen bei Shared Task Mental Model; größter Unterschied: zwischen Pflege und Feuerwehr; Shared Team Mental Model: Hier Überschreitung außerdem von NotSan; größter Unterschied zwischen Notärzt:innen und Berufsfeuerwehr-Beamten:innen.</p> <p>Hierarchie: Followership Trennwert von Notfallsanitäter:innen überschritten; Leadership Trennwert von Notfallsanitäter:innen und Disponent:innen überschritten</p>	<p>Gemeinsam Annahmen der Interviewten: Hohe Arbeitsbelastung der Disponent:innen; KI muss verstehbar sein</p> <p>Einstellung zur KI wird bestimmt durch: Beruflichen Status; KI-Vorwissen; formale Qualifikationen</p> <p>Bei schwacher Ausprägung dieser Faktoren: Hohe Skepsis; KI wird als "Blackbox" wahrgenommen (diffuse, nebulöse Sicht); kaum konkrete Implementierungs-ideen.</p> <p>Positive Einstellung moderiert durch: Reflexion über Arbeits- und Systemkontext; Betrachtung rechtlicher, politischer</p>	<p>Der entwickelte Prozess besteht aus Design, Vorab-Evaluation, Implementierung und Evaluation. Die Evaluation umfasst:</p> <ol style="list-style-type: none"> 1. Aufklärung/Einwilligung 2. Sensorkalibrierung 3. Entspannungsphase 4. Erhebung 5. Abschlussfragebogen <p>Aus Fokusgruppen-interviews resultierende Kernempfehlungen für Mensch-KI-Interaktion:</p> <ol style="list-style-type: none"> 1. Funktionen vorab definieren 2. Systemleistung klären 3. Interaktionszeitpunkte festlegen 4. Kontextrelevante Infos bereitstellen 	<p>Signifikante Unterschiede im Blickverhalten zwischen den Leitstellen.</p> <p>Leitstelle 1: Fokus oben-zentral und rechts-zentral auf Hauptbildschirm;</p> <p>Leitstelle 2: Fokus links-zentral mit Ausbreitung zur Bildschirmmitte; Touchscreens wurden in beiden Setups wenig genutzt; Disponent:innen in Leitstelle 1 berichteten höhere Niveaus mentaler Belastung und Zeitdruck.</p> <p>Eyetracking-Studien sind zeitaufwändig.</p> <p>Die Datenqualität ist mitunter nur mit Einschränkungen bei</p>

			<p>Validierung: Gute bis exzellente Reliabilität und Validität.</p>	<p>und ethischer Aspekte; differenzierte Analyse von Chancen und Risiken</p> <p>Positive Validierung mit unabhängigen Experten.</p>	<ol style="list-style-type: none"> 5. Einfache Korrekturmöglichkeiten 6. Alternativen anbieten 7. KI-Funktionsweise erklären 8. Aus Nutzerverhalten lernen 9. Auswirkungen aufzeigen 10. Anpassungen überschaubar halten <p>Positive Validierung mit Fokusgruppen-Interview.</p>	<p>längeren Aufnahmezeiten in der Leitstelle abzubilden.</p> <p>Validierung: Mittels Überprüfung der korrekten Kalibrierung und des Offsets sowie Bestimmung der Accuracy; überwiegend robuste Datenqualität nach Fallausschluss.</p>
Fazit	<p>Studie zeigt positive Grundvoraussetzungen für technologische Innovationen in Leitstellen; Menschen mit geringer Technikbereitschaft müssen bei der Implementierung berücksichtigt werden; Studie ist begrenzt auf Technikbereitschaft im konkret beschriebenen Setting; Empfehlungen für Folgeforschung: Stichprobengröße erhöhen für bessere Repräsentativität;</p>	<p>Komplexitätsreduzierende Maßnahmen werden dringend benötigt, sofern die komplexitätsbringenden Faktoren nicht auszuschalten sind; technologische Maßnahmen können kontextsensitive DSS mit Nutzer:innen-Feedback und intelligenter Visualisierung sein; weitere Maßnahmen sind auf Ebene der Prozesse (Organisation) und des</p>	<p>Adressatenspezifische Aus- und Fortbildung für HRTs sind wichtig; Fokus auf gemeinsame Zielorientierung, besonders in interprofessionellen Teams; weitere Forschung zu Auswirkungen der Hierarchiewahrnehmung im Rettungsdienst nötig; strukturierte Vorgehensweisen in Kommunikation/Interaktion (z. B. "10 for 10", Übergabeschemata) zur fehlerhandlungssicheren</p>	<p>Problemlage: Hohe mentale Belastung in Leitstellen; KI kann bei komplexen Aufgaben unter Zeitdruck unterstützen; KI muss verstehbar sein; Unterschiede in Überzeugungssystemen.</p> <p>Empfohlene Maßnahmen: Kurzfristig: Verstärkte Diskussionen zu KI-Themen in Leitstellen; praxisnahe Informationsver-</p>	<p>Der Transfer von Mensch-KI-Richtlinien auf Leitstellen lieferte konkrete Design-Elemente für KI-gestützte Notrufbearbeitung; praxisnahe Evaluationsgrundsätze für DSS in Leitstellen konnten beschrieben werden.</p>	<p>Mehr Monitore sind nicht zwangsläufig vorteilhaft; Zwei-Monitor-Setup zeigt Vorteile bei Aufmerksamkeitsverteilung und mentaler Arbeitsbelastung; Empfehlungen für Folgeforschung: Größere, diverse Leitstellen-Setups untersuchen; standardisierte Szenarien entwickeln; kognitive Verarbeitung einbeziehen;</p>

	qualitative Untersuchung der Gründe für geringe Technikbereitschaft/KI-Ablehnung.	Verhaltens anzusiedeln, das vorgeschlagene STOPV-Schema kann dabei helfen; Schichtdauer als Perfomanzfaktor kritisch prüfen.	Kommunikation etablieren; Entwicklung einer "Just Culture".	mittlung. Mittelfristig: KI in Ausbildungsprogramme integrieren; erklärbare und intuitive KI-Systeme entwickeln. Langfristig: Empirische Feldstudien zu KI-Disponent:innen-Interaktion; interdisziplinäre Zusammenarbeit stärken; Theorie an Praxisbeispielen testen.		Beobachtung im Schichtverlauf Drittvariablen berücksichtigen.
Status am 31.10.2024	Veröffentlicht	Im Review			Veröffentlicht	Einreichungsbereit

Anlage 9 – Publikationen

Nachfolgend finden sich die Publikationen 1 bis 6 in aufsteigender Reihenfolge. Falls bereits eingereicht, sind die Publikationen im geforderten Stil des Journals formatiert.