

Universität der Bundeswehr München  
Fakultät für Luft- und Raumfahrttechnik  
Institut für Flugsysteme

# Maschinelle Abschätzung von Situationsbewusstsein zur Umsetzung eines adaptiven Assistenzsystems

Simon Schwerd

Vollständiger Abdruck der von der Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation

Gutachter/Gutachterin:

1. Univ.-Prof. Dr.-Ing. Axel Schulte
2. Univ.-Prof. Dr. Dipl.-Psych. Annette Kluge

Die Dissertation wurde am 08.12.2022 bei der Universität der Bundeswehr München eingereicht und durch die Fakultät für Luft- und Raumfahrttechnik am 15.04.2023 angenommen. Die mündliche Prüfung fand am 21.04.2023 statt.



*Für meinen Vater*



## Vorwort des Autors

Diese Arbeit ist die Zusammenfassung meiner Forschungsarbeit am Institut für Flugsysteme der Universität der Bundeswehr München vom Jahre 2017 bis 2022. Viele Menschen aus meinem Umfeld waren an dieser Arbeit beteiligt. Die wichtigsten Unterstützer möchte ich namentlich erwähnen.

Ein besonderer Dank gilt meinem Doktorvater Herrn Univ.-Prof. Dr.-Ing. Axel Schulte für die Beschäftigung als wissenschaftlicher Mitarbeiter und die Betreuung meiner Doktorarbeit am Institut. Über fünf Jahre lang hat er mich bei meiner Forschung unterstützt und beraten und somit einen wesentlichen Teil zu dem erfolgreichen Abschluss dieser Arbeit beigetragen. Neben der wissenschaftlichen Unterstützung durfte ich zusammen mit ihm großartige Konferenz- und Beschaffungsreisen unternehmen. Ebenso danke ich Dr. Diana Donath für ihre Hilfe in allen Lagen sowie Univ.-Prof. Dr. Dipl.-Psych. Annette Kluge für ihren Beitrag als Zweitgutachterin.

Einen großen Dank richte ich an meine Kollegen aus dem Jet-Simulator. Vom Team rund um diese Versuchsanlage habe ich auf verschiedenste Arten profitiert: Die Umsetzung einer so umfangreichen Simulations- und Aufgabenumgebung wäre mir allein nicht möglich gewesen und wurde maßgeblich von den „Jettis“ Dennis Mund, Felix Heilemann, Sebastian Lindner und Marius Dudek entwickelt. Während der intensiven Zeit der Promotion gab es in diesem Team unzählige leidenschaftliche Diskussionen über technische, wissenschaftliche und gesellschaftliche Themen, in denen ich viel gelernt und gelacht habe! Diese großartige Arbeitsatmosphäre hat mir durch die schwierigen Phasen der Promotion geholfen.

Zuletzt möchte ich mich bei meiner Frau Tanja bedanken, die mich in meiner Promotionszeit auf vielfältige Art unterstützt hat. Sie hat mich in den richtigen Momenten unterstützt, abgelenkt oder beraten und somit auch einen großen Anteil an dieser Arbeit.

Vielen Dank euch und allen, die mich unterstützt haben.

Simon Schwerd

im Mai 2023



## Kurzfassung

Eine effektive Verteilung der Aufmerksamkeit sowie ein hohes Situationsbewusstsein ist wesentlich für die Bedienung moderner Mensch-Maschine-Systeme. Fatale menschliche Fehler im Bereich der Prozesssteuerung, Luftfahrt oder im Autoverkehr können auf niedriges Situationsbewusstsein der Bediener zurückgeführt werden.

Die vorliegende Arbeit beschreibt die Entwicklung und Erprobung eines Assistenzsystems, welches sich an das Situationsbewusstsein des Bedieners eines technischen Systems anpasst. Zentrale Idee des vorgestellten Assistenzsystems ist es, den Bediener auf ihm unbewusste Informationen hinzuweisen.

Hierfür wurde eine maschinelle Abschätzung des Bediener-Situationsbewusstsein entwickelt, welches auf Basis der Analyse von Blickbewegungs- und Interaktionsmessung eine Bewertung darüber erstellt, welche aufgabenrelevanten Informationen einem Bediener nicht bewusst sind. Die maschinelle Bewertung vergleicht hierfür kontinuierlich die vom Bediener wahrgenommenen Informationen mit dem aktuellen Systemzustand. Mithilfe dieses Vergleichs wird ein dynamisches Modell des Situationsbewusstseins erstellt, welches die Wissensbasis des Assistenzsystems repräsentiert. Das Assistenzsystem benachrichtigt auf Basis dieses dynamischen Modells dann den Bediener über aufgabenkritische Informationen, die ihm nicht bewusst sind.

Dieses Konzept wurde für einen Cockpitsimulator implementiert und in drei Mensch-Maschine-Versuchen erprobt. Die drei Studien zielten jeweils auf die Validierung der Analyse-, Bewertungs- und Interventionsfunktion des Assistenzsystems ab. In den Versuchen konnte gezeigt werden, dass das Messsystem in der Lage ist, fehlendes oder falsches Bediener-Situationsbewusstsein informationsspezifisch zu messen. Das Verfahren ist besonders geeignet für Überwachungsaufgaben und besonders präzise bei Informationsänderungen, die vom Bediener nicht vorhersehbar sind. Bewertungsfehler durch das Messsystem traten vor allem dann auf, wenn der Bediener die Änderung von Systemzuständen antizipieren konnte. Die adaptiven Interventionen konnten die Leistung in den Versuchsaufgaben steigern und wurde von Probanden vor allem dann positiv bewertet, wenn die Änderung von Systemzuständen nicht vorhersehbar war.

Im Anschluss an die drei Validierungsstudien, wurde das Assistenzsystem für den Anwendungsfall eines Kampfflugzeugcockpits in einer militärischen Mission mit bemannt-unbemannten Teams umgesetzt. Ein Erprobungsversuch mit Kampfflugzeugpiloten in dynamischen Missionsszenarien zeigte, dass das adaptive Assistenzsystem in einigen Fällen Fehler durch fehlendes Situationsbewusstsein verhindern konnte, jedoch in anderen Situationen störte, da der Pilot an höherpriorisierten Aufgaben arbeitete. Insgesamt ist der gewählte Ansatz für Informationen mit hoher Priorität nützlich, vor allem wenn deren Änderung für einen Bediener nicht vorhersehbar ist.



## Abstract

Effective distribution of attention as well as high situational awareness is essential for the operation of modern human-machine systems. Fatal human errors in the field of process control, aviation or in driving can be attributed to low situational awareness of the operators.

This thesis describes the development and evaluation of an assistance system that adapts to the situation awareness of the operator of a technical system. Based on an online estimation of the operator's situation awareness, targeted notifications are triggered about information, that the operator is not aware of.

For this purpose, a situation awareness model was developed, based on the analysis of eye movement and interaction measurements. It generates an assessment of which task-relevant information an operator is not aware of. The central idea of the model is the continuous comparison between what the operator perceived and the current system state. Based on this model, an intervention strategy for an assistance system is proposed to notify the operator about task-critical information when he is not aware of them. The presented concept was integrated into a fighter simulator and tested in three human-machine trials. The three trials were each aimed at validating the analysis, evaluation, and intervention functions of the assistance system.

In the trials, it was shown that the measurement system can capture missing or wrong operator situation awareness on the perceptual level. The method is particularly suitable for monitoring tasks and is especially accurate for information changes that cannot be predicted by the operator. Measurement errors occurred primarily when the operator could anticipate the changes in system states. Intervention generation was rated positively by subjects for notifications of unanticipated system state change, but less positively when information was of low priority in the situation at hand.

Following the validation trials, the assistance system was implemented for the use case of a combat aircraft cockpit in a military mission with manned-unmanned teams. A test trial with professional fighter pilots in dynamic mission scenarios showed that the adaptive assistance system could prevent errors due to lack of situational awareness in some cases but interfered in other situations as the pilot worked on higher priority tasks. Overall, the selected approach is useful for high-priority information, especially when its change is unpredictable for an operator.



# Inhaltsverzeichnis

1	Einleitung .....	1
2	Stand der Forschung .....	3
2.1	Aufmerksamkeit.....	3
2.2	Situationsbewusstsein .....	17
2.3	Situationsbewusstsein im Kontext menschlicher Fehlleistung.....	30
2.4	Kooperative Automation.....	35
3	Das Situationsbewusstseinsadaptive Assistenzsystem.....	41
3.1	Adaptives Systemverhalten .....	41
3.2	Aufmerksamkeitsanalyse .....	42
3.3	Situationsbewusstseinsbewertung .....	46
3.4	Interventionsgenerierung .....	53
3.5	Gesamtsystemarchitektur.....	54
4	Validierungsstudien .....	60
4.1	Versuchsaufbau.....	60
4.2	Analyse der visuellen Aufmerksamkeit.....	66
4.3	Bewertung des Situationsbewusstseins.....	80
4.4	Interventionsgenerierung .....	102
5	Validierung des Gesamtsystems in einer militärischen Luftoperation .....	120
5.1	Die Arbeitsdomäne: MUM-T in zukünftigen Kampfflugzeugmissionen.....	120
5.2	Gestaltung von SB-Adaptiver Assistenz in MUM-T.....	121
5.3	Validierung der SB-Adaptiven Assistenz .....	129
6	Zusammenfassung & Ausblick .....	142
6.1	Beiträge der Arbeit.....	142
6.2	Verbesserungspotenziale.....	144
6.3	Ausblick .....	145
A.	Abkürzungsverzeichnis.....	146
B.	Literaturverzeichnis .....	147
C.	Abbildungsverzeichnis .....	159
D.	Tabellenverzeichnis.....	163
E.	Anhang .....	164



# 1 EINLEITUNG

Digitalisierung, rechnergestützte Datenverarbeitung und Vernetzung ermöglicht technischen Systemen den Umgang mit einer Informationsmenge, deren Gesamtheit nicht mehr vom Menschen verarbeitet werden kann, sowie eine Automatisierung von kognitiven Aufgaben, deren Ausführung vor ein paar Jahrzehnten noch dem Menschen vorbehalten waren. Diese Entwicklung zeigt sich beispielsweise in selbstfahrenden Autos oder dem Erfolg der Software *AlphaGo* im Wettbewerb mit den besten Spielern der Welt (Silver et al., 2016).

Trotz dieser Erfolge ist der Mensch als Bediener und Überwacher nicht aus technischen Systemen wegzudenken. Zwar ist dessen Aufgabe bei der Bedienung technischer Systeme nur noch selten die manuelle Regelung eines Prozesses, jedoch nimmt er eine überwachende Rolle im Zusammenhang mit den Automatisierungsfunktionen ein (Sheridan, 1992). Jedoch ist die steigende Menge an verarbeiteten Informationen bei gleichzeitiger sinkender Zahl an vom Menschen ausgeführten Aufgaben ist eine Quelle neuer Probleme. Schon in der 1980er-Jahren hat Bainbridge (1983) die „Ironie“ beschrieben, dass die zunehmende Automatisierung dem Bediener nur schwierige bis unmögliche Aufgaben übrig lässt. Hierzu zählt beispielsweise die Überwachung von Systemzuständen zur Feststellung, ob der Bediener intervenieren muss. Für diese Aufgabe ist es elementar, dass der Bediener relevante Systemzustände zum richtigen Zeitpunkt einer Situation beachtet. Doch woher weiß ein Bediener, welche Informationen zu welchem Zeitpunkt relevant sind? Es zeigt sich, dass erfahrene Bediener eine Vorstellung des Systemzustands und der Aufgabensituation haben und auf Basis dieses situativen Modells die relevanten Informationsquellen im Arbeitsplatz auswählen. Gelingt dies, spricht man dem Bediener ein gutes **Situationsbewusstsein** (SB vom engl. *Situation Awareness*) zu. Situationsbewusstsein beschreibt ein kognitives Modell, das im Kontext der mensch-zentrierten Gestaltung von Arbeitsplätzen entwickelt wurde. Es verknüpft die Bedienerentscheidungsfindung und die dynamische Aufgabensituation im Kontext einer spezifischen Arbeitsdomäne. Basierend auf diesem Modell kann beispielsweise bestimmt werden, welche Information in einer bestimmten Aufgabensituation relevant für Bedienerziele und -entscheidungen sind (Endsley & Jones, 2012).

Zusätzliche Automatisierungsfunktionen machen es dem menschlichen Bediener nicht zwingend leichter, ein gutes Situationsbewusstsein aufzubauen. Grund hierfür sind *Out-of-the-loop* Effekte (Endsley & Kiris, 1995) oder eine hohe Komplexität der Automation (Kunze et al., 2019). Aus dieser Problematik resultiert die Frage, wie Automation gestaltet werden kann, um das Zusammenwirken von Automation und dem Bediener zu verbessern. Einen Lösungsansatz liefert hier die **kooperative Automation**, im Englischen auch als *Human-Autonomy-Teaming* bezeichnet, deren zentrale Idee es ist, dass die Automation den Kontext einer Arbeitssituation ausreichend gut versteht, um den Nutzer eigeninitiativ bei seiner Aufgabe zu unterstützen und menschliche Fehler zu verhindern (Schulte, 2013). Folgendes Beispiel soll anschaulich demonstrieren, dass der **kognitive Zustand** des Bedieners eine zentrale Kontextinformation zur Umsetzung kooperativer Automation ist.

*Der Fahrer eines PKWs fährt mithilfe eines Navigationsgeräts durch den Stadtverkehr. Ein Beifahrer sitzt neben ihm und beobachtet den Verkehr. In einem unkritischen Moment meldet das Navigationssystem eine alternative Route und verlangt vom Fahrer eine Bestätigung, damit die neue Route geplant werden kann. Dieser sieht sich die alternative Route an und vergrößert den Kartenausschnitt. Während sich der Fahrer mit der Route beschäftigt, schaltet eine Ampel in 100 Meter auf Rot. Der Beifahrer bemerkt, dass der Fahrer die rote Ampel nicht gesehen hat und warnt den Fahrer, welcher aufblickt und sofort abbremst.*

Der Beifahrer hätte nicht vor der roten Ampel gewarnt, wenn der Fahrer seinen Blick auf die Straße gerichtet hätte. Wie hat der Beifahrer entschieden, dass eine Warnung notwendig ist? Er hat beobachtet, dass die visuelle Aufmerksamkeit des Fahrers auf dem Navigationsgerät liegt. Gleichzeitig hat er eine kritische Änderung der Situation (Ampelsignal) festgestellt und ist zum Schluss gekommen, dass der Fahrer nichts von dieser Änderung weiß und somit ein fehlerhaftes Situationsbewusstsein hat. Folglich warnte der Beifahrer den Fahrer und verhinderte damit, dass es zum Überfahren der Ampel oder einem Auffahrunfall kommt.

Ziel dieser Arbeit ist die Umsetzung und Erprobung eines technischen Assistenzsystems, das eine vergleichbare Art der Kooperation realisiert. Es wird untersucht, wie mit technischen Mitteln festgestellt werden kann, ob ein Bediener adäquates Situationsbewusstsein hat und, ob eine dazu adaptive Warnung nützlich für die Leistung des Mensch-Maschine-Systems ist. Die notwendigen Fähigkeiten des Assistenzsystems lassen sich dabei vom Beispiel der Autofahrt ableiten. Das System muss

- (1) die **Aufmerksamkeit** eines Bedieners kontinuierlich und in Echtzeit **analysieren**,
- (2) daraus das **Situationsbewusstsein** des Bedieners fortwährend **bewerten** und
- (3) auf dieser Basis **intervenieren**.

Im Kapitel 2 wird zunächst der Stand der Forschung im Bereich von Aufmerksamkeit, Situationsbewusstsein und kooperative Automation vorgestellt. Schwerpunkt sind die Messung und Modellierung von SB sowie Studien im Kontext bediener-adaptiver Systeme. Das Kapitel 3 stellt ein technisches Konzept zur Umsetzung des Assistenzsystems vor. Im Zentrum steht die blickbewegungs-basierte Abschätzung von SB und dessen Integration in eine adaptive Strategie, welche bei fehlendem SB den Bediener über kritische Systemänderungen benachrichtigt. Im Anschluss werden in Kapitel 4 drei Validierungsversuche vorgestellt, von denen die ersten beiden zur Validierung des Messsystems und das dritte zur Evaluierung des Gesamtsystems dient. Als Aufgabenumfeld dieser Erprobung dient ein Cockpitsimulator für Kampfflugzeuge des Instituts für Flugsysteme (IFS) der Universität der Bundeswehr München (UniBwM). Der Aufbau der Kapitel 2 - 4 folgt dabei stets den drei Funktionen Aufmerksamkeitsanalyse, Situationsbewusstseinsbewertung und Intervention. Anschließend beschreibt das Kapitel 5 eine Methode zum Transfer des SB-adaptiven Assistenzkonzepts in komplexe Aufgabensituationen am Beispiel einer militärischen Luftoperation für bemannte und unbemannte Kampfflugzeuge. Das letzte Kapitel 6 fasst die erzielten Ergebnisse zusammen und gibt einen Ausblick über mögliche Weiterentwicklungen und Verbesserungen.

## 2 STAND DER FORSCHUNG

Wie eingangs beschrieben, beschäftigt sich die vorliegende Arbeit mit der Analyse von Aufmerksamkeit, der Bewertung von Situationsbewusstsein sowie der Umsetzung einer kooperativen Automation. In diesem Sinne geben die folgenden Abschnitte einen Überblick über den aktuellen Forschungsstand in den für diese Arbeit bedeutsamen Themen **Aufmerksamkeit**, **Situationsbewusstsein** und **kooperative Automation**. Mit Blick auf das Ziel dieser Arbeit wird hierbei ein besonderes Augenmerk auf die arbeitswissenschaftliche sowie messtechnische Perspektive gelegt. Dabei werden einerseits kognitionspsychologische Definitionen und Grundlagen beschrieben und andererseits aktuelle Forschungsstudien diskutiert. Bei Letzterem liegt der Fokus auf Arbeiten, die Anwendung in der Steuerung dynamischer Systeme finden.

### 2.1 AUFMERKSAMKEIT

Die Funktionsweise der menschlichen Aufmerksamkeit ist seit über einem Jahrhundert Gegenstand von empirischen Untersuchungen in der psychologischen Forschung. Aufmerksamkeit steht konzeptionell am Anfang menschlicher Informationsverarbeitung und steht für die kognitive Funktion, die eine Auswahl aus den vielen wahrgenommenen Reizen der Umwelt zur kognitiven Weiterverarbeitung trifft. Dieser Mechanismus ist jedem Menschen aus seiner Erfahrungswirklichkeit vertraut, sodass William James (1890) in den frühen Jahren der Aufmerksamkeitsforschung schrieb:

“EVERYONE KNOWS WHAT ATTENTION IS. IT IS THE TAKING POSSESSION OF MIND IN CLEAR AND VIVID FORM [...] IT IMPLIES WITHDRAWAL FROM SOME THINGS IN ORDER TO DEAL EFFECTIVELY WITH OTHERS.”

Trotz dieser intuitiven Vorstellung von Aufmerksamkeit ist der Mechanismus weder einfach zu definieren noch vollständig ergründet. Dabei kann auch in Frage gestellt werden, ob es überhaupt sinnvoll ist, über das Konstrukt Aufmerksamkeit nachzudenken, wenn es im Grunde unterschiedliche kognitive Funktionen unter sich vereint (Pashler, 1998, S. 226).

Eine weit verbreitete Metapher für Aufmerksamkeit ist die einer Taschenlampe, die ausgewählte Bereiche im Inneren des Menschen und der Umwelt beleuchtet, wobei Beleuchtung bildlich für die Auswahl dieser Information zur Verarbeitung steht. Ein breiter Lichtstrahl steht für eine verteilte Auswahl und damit für verteilte Aufmerksamkeit (engl. *divided attention*), während ein eng fokussierter Lichtstrahl die selektive und fokussierte Aufmerksamkeit (engl. *selective attention*) darstellt. Eine alternative Interpretation der Lichtstrahlbreite ist, dass die selektive Aufmerksamkeit breit oder fokussiert sein kann. Die Batterie der Taschenlampe steht für die Begrenzung von anhaltender Aufmerksamkeit (engl. *sustained attention*).

Erfahrungen im zweiten Weltkrieg führten zu den ersten systematischen Untersuchungen von Aufmerksamkeit. Dabei fiel auf, dass Radarbediener die auf sie eintreffenden Informationen und Signale nicht vollumfänglich verarbeiten konnten, beispielsweise wenn sie kurz aufeinander folgten (Welford, 1952). Die zentrale Frage der Aufmerksamkeitsforschung war zunächst, wo dieser kognitive Verarbeitungseingpass (engl. *bottleneck*) liegt und wie der Filtermechanismus funktioniert, der eintreffende Reize zur Verarbeitung auswählt und andere ignoriert. Etwas später kam dann auch die Frage auf, inwiefern eine parallele Verarbeitung mehrerer Informationsquellen möglich ist, was die Theorien über Aufmerksamkeit als Ressource hervorbrachte (Egeth & Kahneman, 1975; Knowles, 1963).

Diese zwei Gestalten von Aufmerksamkeit sind weiterhin etablierter Bestandteil von Modellen menschlicher Aufmerksamkeit. Aufmerksamkeit wird zum einen als Ressource verstanden, die bei der Verarbeitung von Informationen notwendig ist, und zum anderen als Selektionsmechanismus, der unter den vielen Informationen der Arbeitsumgebung diejenigen auswählt, die verarbeitet werden (Wickens, 2021).

Die folgenden Abschnitte beschreiben Modelle von Aufmerksamkeit und deren Relevanz für die Mensch-Maschine-Interaktion. Dabei konzentriert sich diese Arbeit besonders auf die Gestalt der selektiven Aufmerksamkeit. Eine vollumfängliche Darstellung des Forschungsstands zur menschlichen Aufmerksamkeit würde den Rahmen dieser Arbeit übersteigen. Hierfür wird auf die Quellen verwiesen, die als Basis für das folgende Kapitel dienen (Oberauer, 2019; Pashler, 1998; Styles, 2005; Wickens, 2021; Wickens et al., 2022).

### 2.1.1 DIE MENSCHLICHE INFORMATIONSVERARBEITUNG

Ein hilfreiches Mittel zur Beschreibung von Aufmerksamkeit ist das Stufenmodell der menschlichen Informationsverarbeitung, dargestellt in Abbildung 2-1, welches von Remington und Loft (2015) sowie Wickens et al. (2022) adaptiert wurde. Gezeigt sind die verschiedenen Stufen der Verarbeitung und der Fluss von Information, wenn Menschen zielgerichtet eine Aufgabe verfolgen.

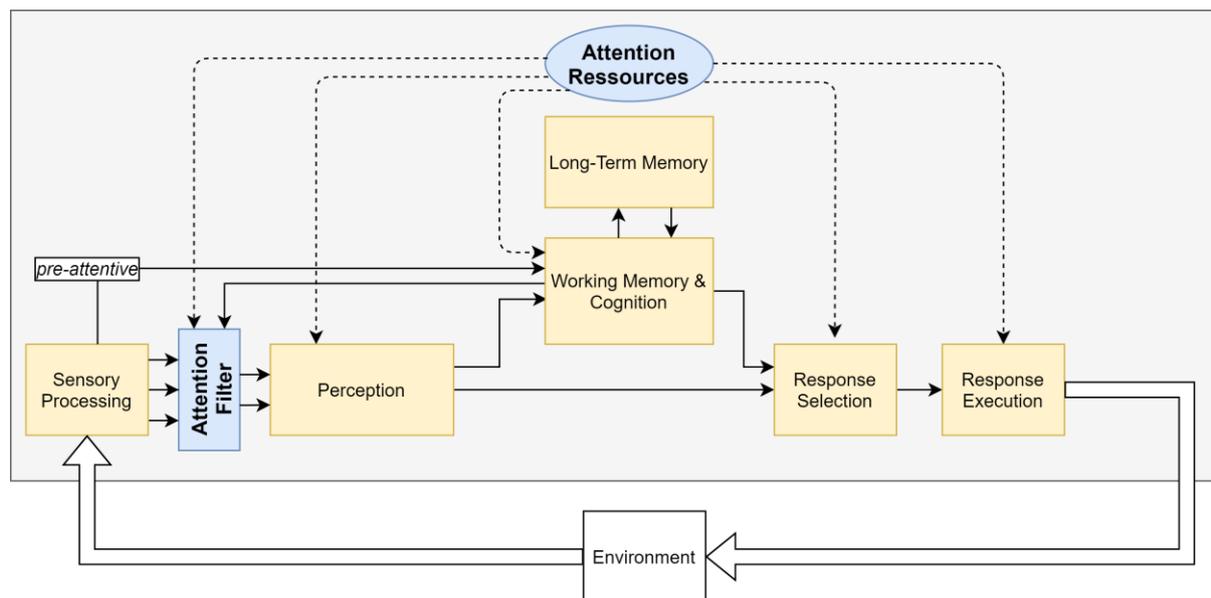


Abbildung 2-1 - Ein Model der menschlichen Informationsverarbeitung nach (Remington & Loft, 2015; Wickens et al., 2022)

Zu Beginn der Informationsverarbeitung steht die sensorische Verarbeitung von Reizen durch die menschlichen Sinne wie die von der Netzhaut des Auges aufgenommenen Reize durch einen Blick auf eine Anzeigegerät. Die sensorische Aufnahme bedeutet nicht, dass die Information auch weiterverarbeitet wird. Die Auswahl der zu verarbeitenden Reize trifft ein Filter der selektiven Aufmerksamkeit, der nicht alle aufgenommenen Reize weiterleitet, sondern im Idealfall nach relevanten Informationen filtert. In Arbeitssituationen wird dieser Selektionsmechanismus gesteuert durch die aktuelle Aufgabe und Zielsetzung des Menschen (Remington & Loft, 2015). Dieser zielorientierte Aufmerksamkeitsmodus wird in der Literatur oft als *top-down attention* bezeichnet. Der zielgesteuerten Aufmerksamkeit steht das „Einfangen“ von Aufmerksamkeit durch einen salienten Reiz gegenüber, auch als *bottom-up attention* beschrieben. Ein Reiz ist salient, wenn er hervorstechende Charakteristiken hat, die vor selektiver Filterung (engl. *pre-attentive*) bereits

verarbeitet werden. Beispiele beim Umgang mit technischen Systemen sind ein lautes Warnsignal oder ein rot blinkendes Symbol.

Die selektierten Informationen werden wahrgenommen und gelangen in das Arbeitsgedächtnis, was bedeutet, dass sie dem Menschen bewusst werden und gegebenenfalls eine kognitive Verarbeitung folgt. Kognitive Verarbeitung bezeichnet in diesem Modell geheimhin das „Nachdenken“, Einordnen und Transformieren von Informationen im Arbeitsgedächtnis. Nach der Verarbeitung steht die Entscheidung für eine Reaktion und die Durchführung derselben, welche wiederum einen Einfluss auf die Umwelt hat. Neben den gefilterten Informationen werden auch ignorierte Reize bis zu einem gewissen Grad verarbeitet (*pre-attentive*), was beispielsweise zu einer plötzlichen Neuausrichtung der Aufmerksamkeit bei salienten Reizen führen kann. In dieser Modellvorstellung werden alle Prozesse durch eine Aufmerksamkeitsfunktion mit kognitiven Ressourcen versorgt. Die Menge an mentalen Ressourcen ist jedoch limitiert, wodurch Leistungsdekremente bei einer zu hohen Aufgabenbelastungen entstehen. Beispielsweise verlangen geübte Aufgaben einen niedrigeren kognitiven Aufwand als Tätigkeiten, die dem Menschen komplett neu sind (Schneider & Shiffrin, 1977). Zusammenfassend zeigt sich Aufmerksamkeit also in zwei Ausprägungen, die beide zu einem gewissen Grad vom Menschen gesteuert werden (Wickens, 2021; Wickens et al., 2022):

- **Selektion von Informationen aus den sensorisch verarbeiteten Reizen:** Der erste Mechanismus, die Selektion von Informationen, bezieht sich auf die Auswahl von relevanten Stimuli unter allen eintreffenden Reizen. Dieser Mechanismus verhindert die Verarbeitung von unwichtigen Reizen zugunsten aufgaben-relevanter Informationen.
- **Verteilung der verfügbaren Aufmerksamkeitsressourcen:** Der zweite Mechanismus verteilt kognitive Ressourcen auf Verarbeitung ausgewählter Informationen, Auswahl und Durchführung einer Antwort. Dieser Mechanismus ist in Arbeitssituationen besonders relevant, in denen ein Mensch die Bearbeitung konkurrierender Aufgaben verfolgt (engl. *Multi-Tasking*).

Grundsätzlich weisen die Autoren verschiedener Aufmerksamkeitsmodelle darauf hin, dass kein Modell bisher die volle Bandbreite aller empirischen Daten erklären kann (Oberauer, 2019; Pashler, 1998; Wickens, 2021). Aus Sicht der Kognitionsergonomie können diese Theorien aber als Grundlage für nützliche Prinzipien und Modelle zur Entwicklung von Mensch-Maschine-Systemen dienen (Wickens, 2021). Da im weiteren Verlauf dieser Arbeit ein System zur Messung von visueller selektiver Aufmerksamkeit entwickelt wird, wird im Folgenden die selektive Aufmerksamkeit genauer beleuchtet.

### 2.1.2 AUFMERKSAMKEIT ALS SELEKTIONSMECHANISMUS

Von selektiver Aufmerksamkeit wird gemeinhin gesprochen, wenn sich Menschen auf eine aufgabenrelevante Informationsquelle konzentrieren, während sie andere Reize ausblenden. In einer Arbeitssituation setzt eine Art Aufmerksamkeitskontrolle (engl. *attentional control*) den Filter auf aufgabenrelevante Reize, welche vorzugsweise verarbeitet werden. Wie genau dieser Filter funktioniert und in welchem Maße unbeachtete Reize verarbeitet werden, war eine der zentralen Forschungsfragen der frühen Aufmerksamkeitsstudien.

## Früher Filter versus später Filter

Im Allgemeinen liegt jeder Filtertheorie die Vorstellung zugrunde, dass vor dem selektiven Filter eine parallele und unbewusste Wahrnehmung von Reizen stattfindet, während nach der Selektion eine serielle und bewusste Verarbeitung durchgeführt wird. Diese Vorstellung basiert auf der empirischen Beobachtung, dass es einen zentralen Engpass in der menschlichen Informationsverarbeitung gibt. Das zeigt sich beispielsweise durch die psychologische Refraktärperiode, bei dem die Reaktion eines Probanden auf einen zweiten von zwei schnell aufeinander folgenden Reizen verzögert ist (Welford, 1952). Um dieser Limitierung Rechnung zu tragen, muss es folglich eine Art selektiven Mechanismus geben, der verhindert, dass alle Reize der Umwelt ungehindert auf diesen Engpass treffen. Die ersten Theorien in diesem Gebiet gingen zunächst von einer frühen Filterung (engl. *early-filter-theory*) von Reizen auf Basis ihrer physikalischen Attribute aus, wie beispielsweise der Tonlage eines Sprechers beim bekannten Cocktail-Party-Effekt (Broadbent, 1958). In darauffolgenden Studien ergaben sich widersprüchliche Experimentalergebnisse, die für eine parallele Verarbeitung aller Reize bis zu einer semantischen Ebene sprachen, wobei nur das wichtigste Signal in das Bewusstsein dringt und verarbeitet wird. Die daraus resultierende Theorie einer späten Filterung (engl. *late filter theory*) besagte, dass relevante und irrelevante Stimuli ohne Limitierung parallel verarbeitet werden können und der Engpass erst in den Gedächtnis- und Verhaltensfunktionen liegt (Duncan, 1980). Beispielsweise bemerkten Versuchsteilnehmer, die sich auf einen von zwei Sprechern konzentrieren sollten, wenn der nicht zu beachtende Sprecher ihren Namen sagte (Moray, 1959). Studien im Bereich der visuellen selektiven Aufmerksamkeit bestätigten, dass sowohl semantische als auch kategorische Attribute von irrelevanten Reizen durchaus verarbeitet werden, aber nicht unbedingt in das Bewusstsein von Versuchsteilnehmern dringen. Beispielsweise sank die Antwortgeschwindigkeit bei Zielsuchaufgaben, wenn ein Ablenkungsstimulus vorhanden war, dessen Identität mit einer nicht zum Ziel passenden Antwort verbunden war (Eriksen & Eriksen, 1974). Ein anderes Experiment von Moray (1970) zeigte, dass Versuchsteilnehmer physiologische Reaktionen auf Wörter in nicht beachteten auditiven Erzählungen zeigten, welche im Vorgang mit einem Elektroschock assoziiert wurden. Die klassische Filtertheorie, die von einer strikten seriellen Verarbeitung nach einem frühen Filter ausging, verlor damit Plausibilität und wurde stetig an neue empirische Ergebnisse angepasst.

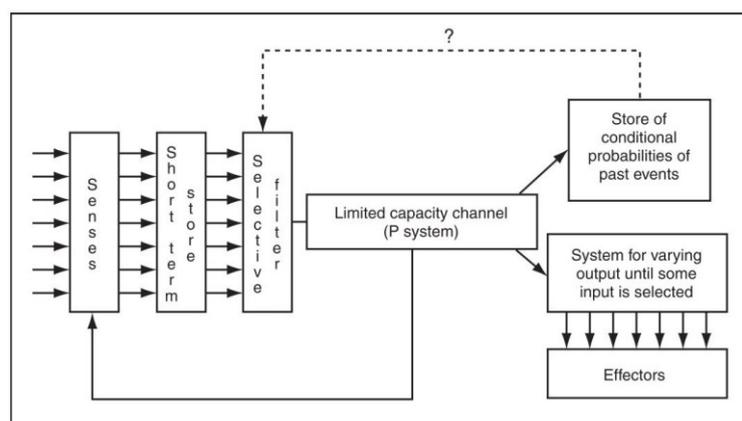


Abbildung 2-2 - Filtertheorie nach Broadbent (Styles, 2005, S. 21)

Im Streit über die Theorien von früher oder später Filterung argumentierten einige Wissenschaftler, dass die widersprüchlichen Ergebnisse durch unterschiedliche Experimentalparadigmen zustande gekommen sind, welche entweder die eine oder die andere Version der Filterung darstellten. Diese Debatte mündete in einem hybriden Modell der selektiven Aufmerksamkeit, das postulierte, dass der

Filter abhängig von den Aufgabenanforderungen ist, und welches mittlerweile als die Belastungstheorie der selektiven Aufmerksamkeit bekannt ist (Yantis & Johnston, 1990).

## Belastungstheorie

Die Belastungstheorie (engl. *load theory*) besagt, dass die selektive Aufmerksamkeit abhängig ist von der Schwierigkeit einer Aufgabe in den Dimensionen Wahrnehmung, Arbeitsgedächtnis und notwendige Kontrolle zur Ausführung der Aufgabe (engl. *cognitive control*) (Lavie et al., 2004). Das Modell besagt, dass die Wahrnehmung ein ressourcen-limitierter Prozess ist, der automatisiert abläuft, bis die Ressourcen erschöpft sind. Bei hoher Wahrnehmungsbelastung (engl. *perceptual load*) erfolgt eine frühe Filterung von ablenkenden Stimuli, während bei niedriger Wahrnehmungsbelastung eine späte Filterung und damit eine tiefergehende Verarbeitung von ablenkenden Stimuli erfolgt. Wahrnehmungsbelastung kann beispielsweise steigen durch eine größere Anzahl von Stimuli oder schwierigere Erkennungskriterien von relevanten Stimuli (bspw. ein Wort gegenüber einem Buchstaben). Experimente zeigen, dass die Erhöhung dieser Belastung dazu führt, dass sich Antwortzeiten bei Zielsuchaufgaben mit Ablenkungsreizen sogar verbessern, was darauf zurückgeführt wird, dass ablenkende Reize nicht mehr verarbeitet werden, weil die Wahrnehmungskapazität erschöpft ist. Hohe Wahrnehmungsbelastung fördert damit auch das Phänomen der Aufmerksamkeitsblindheit (engl. *inattentional blindness*). Eine zweite Aussage der Belastungstheorie ist, dass eine hohe Belastung von Prozessen der kognitiven Kontrolle (bspw. das Arbeitsgedächtnis) die Fähigkeit reduziert, irrelevante Stimuli zu ignorieren. Dies führt zu steigenden Reaktionszeiten durch ablenkende Stimuli in Experimenten mit hoher kognitiver Belastung bei gleichzeitiger niedriger Belastung der Wahrnehmung. Eine Vielzahl empirischer Ergebnisse aus Verhaltensstudien sowie neurologischer Arbeiten sprechen für die Belastungstheorie. Dabei ist der Zusammenhang zwischen Wahrnehmungsbelastung und Aufmerksamkeitsblindheit besser nachgewiesen worden, als die Auswirkung hoher kognitiver Belastung auf die selektive Aufmerksamkeit (Matias et al., 2021). In diesem Kontext wird auch über eine genaue Operationalisierung von Belastung für Kognition und Wahrnehmung sowie deren Interaktion diskutiert (Murphy et al., 2016).

### 2.1.3 SELEKTIVE AUFMERKSAMKEIT IN DER MENSCH-MASCHINE-INTERAKTION

An dieser Stelle soll der Blick nun in Richtung der Mensch-Maschine-Interaktion gerichtet werden und die Frage, inwiefern selektive Aufmerksamkeit für die Leistungsfähigkeit von Menschen bei der Bedienung technischer Systeme eine Rolle spielt. Durch den Einsatz softwarebasierter Anzeigesysteme in modernen Arbeitsplätzen machen visuelle Aufgaben den Großteil der Aufgaben aus, weshalb der Fokus in dieser Arbeit auch auf selektiver **visueller** Aufmerksamkeit liegt. Hierfür ist es nützlich zu definieren, in welchen Aufgabentypen selektive Aufmerksamkeit eine Rolle spielt (Wickens et al., 2022):

- Überwachende Kontrolle von Systemwerten und -zuständen (engl. *Supervisory Control*)
- Bemerkens von unerwarteten Informationen (engl. *Noticing*)
- Suchen nach Zielen mit definierten Kriterien
- Lesen von Texten

Viele typische Arbeitssituationen beinhalten eine Kombination dieser Aufgaben, wie beispielsweise bei einem Piloten, der in regelmäßigen Abständen seine Flughöhe und -geschwindigkeit überwacht,

dann eine Warnung des Kollisionswarnsystems bemerkt und letztlich im Luftraum vor sich nach dem Flugzeug sucht, das eine Kollisionsgefahr darstellt. Die folgenden Abschnitte beschreiben, welche Rolle die selektive Aufmerksamkeit im Zusammenhang mit diesen vier Aufgabentypen spielt.

## Überwachende Kontrolle

Der Begriff der überwachenden Kontrolle beschreibt den Kontrollmodus eines menschlichen Bedieners, der ein hochautomatisiertes, technisches System steuert. In diesem Modus nimmt der Bediener eine überwachende Rolle ein, in der er überwiegend die Systemzustände beobachtet (engl. *monitoring*) und sofern notwendig durch diskrete Befehle anpasst (Sheridan, 1992). Der tatsächliche Prozess wird mittels Automatisierung von einem Rechner gesteuert. Dieser Modus unterscheidet sich von der manuell-kontinuierlichen Kontrolle eines technischen Prozesses beispielsweise über einen Joystick.

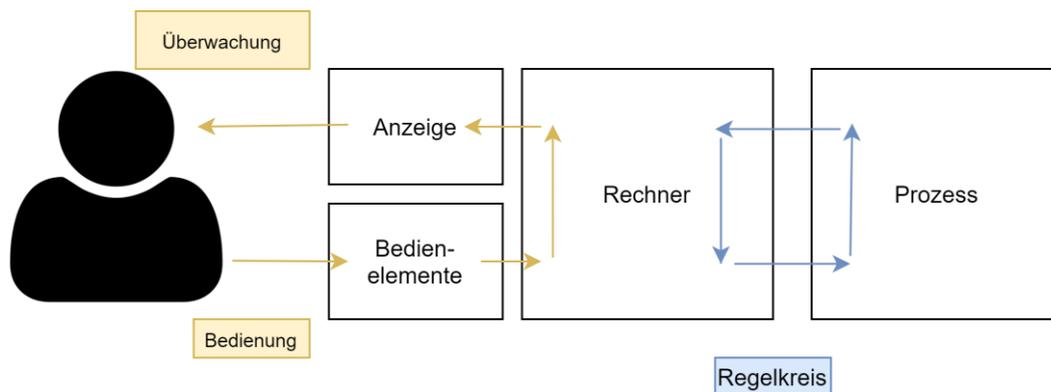


Abbildung 2-3 - Überwachende Kontrolle eines technischen Prozesses nach Sheridan (1992)

In diesem Kontrollmodus ist die zielbasierte selektive Aufmerksamkeit wesentlich. Je nach Arbeitsumfeld und Aufgabe, richtet ein Bediener seine Aufmerksamkeit sequenziell auf verschiedene Informationsquellen und überprüft dort, ob sich die Systemparameter noch im gewünschten Bereich befinden. Beispielsweise überwacht ein Pilot beim Flug mit Autopiloten kontinuierlich Flugparameter und überprüft dabei, ob sie der gewünschten Flugtrajektorie entsprechen. Mit zunehmender Automatisierung nimmt die Aufgabe der Überwachung eine größere Rolle ein. Dies ist problematisch, da der Bediener nicht direkt im Kontrollregelkreis eingebunden ist und nur in Fehlerfällen eingreifen soll. Die daraus resultierenden negativen Effekte werden von Bainbridge (1983) als „Ironien der Automation“ bezeichnet.

Nach dem SEEV-Modell von Wickens et al. (2007) wird die Aufmerksamkeit in diesem Arbeitsmodus von vier Attributen einer Informationsquelle beeinflusst: Salienz, Aufwand, Erwartung und Nutzen. Dabei stellen Salienz und Aufwand die Einflüsse der Arbeitsumgebung dar, während Erwartung und Nutzen aus dem Wissen des Bedieners resultieren. Mithilfe einer Modellierung dieser vier Parameter konnten die Überwachungsmuster von menschlichen Bedienern in verschiedenen Disziplinen vorhergesagt werden (Grundgeiger et al., 2021; Steelman-Allen & McCarley, 2011).

## Bemerken von unerwarteten Informationen

Die Tendenz eines Bedieners bei der überwachenden Kontrolle nur aufgabenorientierte Informationen zu verarbeiten, birgt die Gefahr, dass unerwartete Informationen nicht bemerkt werden. Wenn dies geschieht, wird von Unaufmerksamkeitsblindheit gesprochen (Mack, 2003) und tritt vor allem dann auf, wenn eine sichtbare Änderung nicht erwartet wird. Simons und Chabris (1999) führten in diesem Kontext ein berühmtes Experiment durch, in der Probanden zählen sollten, wie oft

im Video eines Basketballspiels der Ball zwischen weiß oder schwarz gekleideten Spielern gepasst wird. Während dieser visuellen Aufgabe lief ein Mann in einem schwarzen Gorillakostüm durch das Video. Über die Hälfte der Probanden bemerkten den Mann im Gorillakostüm nicht, wobei dieser Effekt besonders ausgeprägt war, wenn sich die Teilnehmer auf Spieler mit weißem T-Shirt konzentrieren mussten. Die Autoren erklärten sich das damit, dass unerwartete Reize im Sichtfeld eher erkannt werden, wenn sie visuelle Eigenschaften mit aufgabenrelevanten Reizen teilten. Unaufmerksamkeitsblindheit ist somit eine Limitation der menschlichen Verarbeitung visueller Reize und resultiert letztendlich aus dem Filter der selektiven Aufmerksamkeit, wodurch nur aufgabenrelevante Reize (festgelegt durch die Erwartung des Menschen) verarbeitet. Im Kontext der Luftfahrt domäne fand Shorrock (2007) in Unfallanalysen heraus, dass ein Drittel aller Fehler von Fluglotsen mit Problemen in der Wahrnehmung von relevanten Informationen zusammenhingen, selbst wenn diese relevanten Informationsquellen zuvor mit einem Blick fixiert wurden.

Aufbauend auf dem SEEV Modell, erweiterte Wickens (2015) das SEEV Modell um eine *Noticing*-Komponente, die Vorhersagen darüber trifft, wie wahrscheinlich das Bemerkten einer Meldung am Arbeitsplatz ist. Dabei wird die *Noticing*-Komponente definiert über Salienz, Erwartung und Exzentrizität. Exzentrizität wird dabei zum einen durch den Winkelabstand von Interessensgebieten zur Warnung und zum anderen durch die kognitive Beanspruchung durch Aufgaben modelliert. Bei hoher kognitiver Beanspruchung wird von einer fokussierten Aufmerksamkeit ausgegangen, die die Exzentrizität erhöht. In einem Versuch, bei dem die Teilnehmer Tetris spielten und gleichzeitig eine Anzeige überwachen musste, konnte das NSEEV-Modell zwei Leistungsmaße sehr genau vorhersagen: wie oft Systemwarnungen nicht bemerkt werden und zum anderen, falls die Warnung bemerkt wird, wie schnell auf Warnungen reagiert wird.

### Suchen nach Zielen mit definierten Kriterien

Ein dritter Aufgabentyp in Arbeitssituationen ist das Suchen nach Zielen. Beispielsweise sucht ein Kampfflugzeugpilot nach einem anderen Flugzeug in seiner Umgebung oder einem Objekt auf der taktischen Karte. Gegenüber dem „Bemerkten“ unterscheidet sich das Suchen dahingehend, dass das Zielobjekt persistent an einem Ort im Raum liegt, als dass es ein diskretes Ereignis im zeitlichen Verlauf ist. Die menschliche Leistung bei der Suche ist limitiert, da einerseits nicht alles im Blickfeld gleichzeitig wahrgenommen und verarbeitet werden kann und andererseits das Blickfeld in vielen Arbeitsplätzen durch Augen-, Kopf- und Körperbewegungen neu ausgerichtet werden. Die meiste Forschung in diesem Gebiet bezieht sich jedoch auf das Suchen innerhalb des Blickfelds. Die Suchgeschwindigkeit ist abhängig von zwei zentralen Faktoren: Eigenschaften der Suchaufgabe, wie Anzahl der Objekte, ablenkende Stimuli und *Clutter*, sowie der Unterscheidbarkeit der Objekte untereinander. Das Suchverhalten des Menschen wird von fünf verschiedenen Faktoren beeinflusst (Wolfe & Horowitz, 2017):

- Salienz der Ziele: Objekte, die sich von ihrer Umgebung unterscheiden ziehen Aufmerksamkeit an.
- A-priori bekannte Merkmale des Zielobjekts: Wenn das Zielobjekt durch eine einzelne, saliente Eigenschaft von allen Nicht-Zielobjekten unterscheidbar ist, findet die visuelle Suche annähernd parallel statt. Dieser sogenannte *Pop-out*-Effekt wird verstärkt, wenn die Nicht-Zielobjekte sehr homogen sind.
- Struktur der Suchszenen: Vor allem beim Suchen in naturalistischen Szenen leitet auch Plausibilität die Verteilung der Aufmerksamkeit.

- Erfahrung: Durch Erfahrung in der Suchaufgabe wird die Verteilung der Aufmerksamkeit beeinflusst.
- Wertigkeit der Ziele: Wenn bestimmte Zielobjekteigenschaften (z.B. Farbe) mit einer höheren Wertigkeit im Sinne der Aufgabe assoziiert sind, wird die Aufmerksamkeit bei Suchaufgaben bevorzugt auf diese Objekte gerichtet.

## Lesen

Das Wahrnehmen und Verstehen von Text und Symbolen ist ein zentraler Aspekt in der Bedienung von Mensch-Maschine-Systemen. Die kognitive Verarbeitung ist dabei hierarchisch aufgebaut von der Wahrnehmung visueller Merkmale (z.B. Linien, Bögen), über Buchstaben und Wörter bis hin zu Sätzen. Verschiedene Studien sprechen dafür, dass bei geübten Lesern auf der Ebene der Merkmale, Buchstaben bis hin zur Wörferebene eine parallele Verarbeitung stattfindet und diese durch eine Mustererkennung als Einheit wahrgenommen werden. So zeigte Laberge (1973), dass bekannte Buchstaben wie „b“ oder „d“ verarbeitet werden, selbst wenn diese nicht relevant für die aktuelle Aufgabe sind. Dasselbe traf nicht für „künstliche“ Buchstaben zu, die aus den gleichen Merkmalen bestanden. Diese effiziente Verarbeitung hängt also mit den Erfahrungen im Lesen zusammen, welche zu einem kognitiven Automatismus der Verarbeitung führt. Dies konnte auch auf Ebene von hochfrequenten Wörtern im Englischen wie beispielsweise „the“ nachgewiesen werden (Haber & Schindler, 1981). Dabei wurde die Leistung von Versuchsteilnehmern im Korrekturlesen untersucht und dabei festgestellt, dass Fehler in frequenten gegenüber nicht-frequenten Wörtern weniger oft gefunden werden, sofern der Fehler nicht die Form des Wortes verändert. Neben der Verarbeitung von Merkmalen, Buchstaben und Wörtern spielen auch Erwartung und Kontext eine Rolle für das Textverständnis<sup>1</sup>. Durch Kontext und Syntax ermöglicht Sprache die Vorhersage von Merkmalen, Buchstaben und Wörtern ermöglichen. Wenn diese Redundanz einer Sprache reduziert wird, sinkt auch der Anteil der kontextgetriebenen Verarbeitung von Text (Tulving et al., 1964). Das Phänomen des Automatismus ist auch ein entscheidender Aspekt für den Leistungsunterschied zwischen erfahrenden und unerfahrenen Bedienern wie beispielsweise bei Piloten nachgewiesen wurde. In der Studie von wurde beispielsweise von Lounis et al. (2021) nachgewiesen, dass erfahrene Piloten relevante Anzeigen kürzer fixieren als Novizen.

### 2.1.4 MESSUNG SELEKTIVER AUFMERKSAMKEIT

Wie das vorangegangene Kapitel beschreibt, handelt es sich bei selektiver Aufmerksamkeit um eine kognitive Funktion, die sich der direkten Beobachtung von außen entzieht. Um dennoch Aussagen über die Natur dieses Konstrukts treffen zu können, muss Aufmerksamkeit operationalisiert werden. Operationalisierung beschreibt in der Psychologie den Prozess, nicht beobachtbaren Konstrukten beobachtbare Größen zuzuordnen und sie damit bedeutsam quantifizierbar zu machen. Für Aufmerksamkeit gibt es hierfür drei Ansätze: die Messung von **Leistungsmaßen**, **physiologischen Parametern** des Bedieners oder des **beobachtbaren menschlichen Verhaltens** im Umgang mit dem System. Aufbauend auf diesen Messungen gibt es Metriken, denen das Konstrukt Aufmerksamkeit zugeordnet wird.

#### 2.1.4.1 Leistungsmaße

Grundlegende Experimente im Kontext der Aufmerksamkeitsforschung bedienen sich oft an der Variation von Aufgabentypen, in denen Aufmerksamkeit mithilfe eines Leistungsmaßes

---

<sup>1</sup> Beispielsweise fügte das Gehirn des Autors beim Korrekturlesen dieser Arbeit Wörter in Sätze, die dort gar nicht standen.

operationalisiert wird. Zu klassischen Experimentalaufgaben zählen beispielsweise Filteraufgaben, bei denen der Versuchsteilnehmer mit mehreren Stimuli konfrontiert wird und eine vordefinierte Eigenschaft eines Stimulus mit einem bestimmten Selektionskriterium finden und melden muss. Ein bekannter Vertreter dieses Aufgabentyps ist der sogenannte Flanker-Task (Eriksen & Eriksen, 1974). Hierbei können Fehlerraten und Reaktionszeit als Messgröße für Aufmerksamkeit dienen. Ein weiterer Aufgabentyp sind die Überwachungs- und Suchaufgaben, bei denen sich die gestellte Aufgabe meist auf ein kategorisches Attribut eines von vielen Stimuli bezieht und sich die Aufgabe über einen Zeitraum erstreckt (Pashler, 1998, S. 29). Hierbei bieten sich auch Fehlerraten und Reaktionszeiten als Operationalisierung von Aufmerksamkeit an. Zuletzt gibt es auch Aufgabentypen, die Versuchsteilnehmern nur einen einzigen Stimulus präsentieren und nur die Anwesenheit eines bestimmten Attributes im Stimulus gemeldet werden soll. Diese Experimente werden oft im Kontext der Signalerkennungstheorie benutzt (Tanner & Swets, 1954).

#### 2.1.4.2 Physiologische Messverfahren

Die meisten physiologischen Verfahren konzentrieren sich auf die Messung von Signalen des menschlichen Gehirns. Andere physiologische Parameter, wie beispielsweise die Herzrate, korrelieren mit Zuständen wie Müdigkeit (Patel et al., 2011) oder Stress (H. Kim et al., 2018), welche sich negativ auf die Aufmerksamkeit eines Bedieners auswirken. Sie werden jedoch selten zur direkten Bestimmung von Aufmerksamkeit verwendet. Ein Grund hierfür ist sicher auch, dass Aufmerksamkeit als kognitives Phänomen verstanden wird, dessen zeitliche Dynamik wesentlich höher ist als beispielsweise die physiologische Reaktion gemessen an der Herzrate (Mund, 2023). Im Folgenden werden zwei der etablierten physiologischen Messverfahren zur Messung von Aufmerksamkeit aufgeführt.

##### Funktionelle Nahinfrarotspektroskopie (fNIRS):

Die fNIRS ist ein nicht-invasives, bildgebendes Verfahren, das die Hirnaktivität durch Änderungen der optischen Eigenschaften des Hirngewebes quantifiziert. Mittels Nahinfrarotspektroskopie wird dabei die regionale Sauerstoffkonzentration im Hirngewebe gemessen, die im Zusammenhang mit der kognitiven Aktivität steht (Ferrari & Quaresima, 2012). Dieses Verfahren wurde beispielsweise genutzt, um in einer Wachsamkeitsaufgabe zu erkennen, ob ein Versuchsteilnehmer seine Aufmerksamkeit auf eigene, zerstreute Gedanken richtet (Bogler et al., 2014; Durantin et al., 2015). Die Autoren anderer Studien setzen fNIRS ein, um den Grad der mentalen Beanspruchung zu messen (Herff et al., 2013). Obwohl Aufmerksamkeit mit der Beanspruchung zusammenhängt, wird hier aber nicht direkt das Konstrukt Aufmerksamkeit operationalisiert. Beispielsweise konnte mit fNIRS erfolgreich klassifiziert werden, ob ein Pilot manuell oder automatisiert (Verdière et al., 2018). Die Autoren begründeten das mit einem unterschiedlichen Grad an „Involviertheit“ des Piloten in beiden Fällen.

##### Elektroenzephalographie (EEG):

EEG misst die Änderungen des elektrischen Potentials des äußeren Gehirns, welche meist nicht-invasiv mittels Elektroden auf der Kopfhaut gemessen werden. Das elektrische Potential entsteht durch synchrone Aktivität kortikaler Neuronen und repräsentiert somit die Gehirnaktivitäten in den äußeren Gehirnregionen (Jackson & Bolger, 2014). Dieses Messverfahren wurde in einigen Studien zur Detektion von auditiver Aufmerksamkeit verwendet (Belo et al., 2021). Im Zusammenhang mit Luftfahrtanwendungen konnte beispielsweise mit einer Genauigkeit von 72% vorhergesagt werden, ob ein auditiver Alarm wahrgenommen und verarbeitet wurde (Dehais et al., 2019). Eine andere Studie nutzte EEG, um Unsicherheiten in einem kognitiven Modell zu verbessern, das nicht

wahrgenommene Alarmsignale im Flugzeugcockpit vorhersagen sollte (Klaproth et al., 2020). Aber auch im Zusammenhang mit Wachsamkeit in visuellen Aufgaben wurde EEG erfolgreich eingesetzt, um fehlende Aufmerksamkeit vorherzusagen (Martel et al., 2014; Sauvet et al., 2014).

### 2.1.4.3 Bedienungsbeobachtung

Die bisher genannten Messverfahren sind mit moderater Genauigkeit in der Lage durch Messungen der Gehirnaktivität zwischen Aufmerksamkeit und Unaufmerksamkeit (engl. *inattention blindness*) zu unterscheiden (Mack, 2003). Gegenüber diesen physiologischen Messverfahren werden bei der Beobachtung von Bedienerverhalten nicht Reaktionen des vegetativen Nervensystems gemessen, sondern Metriken der Interaktion des Bedieners mit dem technischen System. Diese Messverfahren besitzen damit einen Bezug zum Aufgabenkontext des Bedieners, was sie von physiologischen Messverfahren abgrenzt. Beispielsweise kann eine Klickrate auf ein Kontrollelement nur bei der Bedienung eines entsprechenden Systems gemessen werden, während die Herzrate auch ohne diesen Aufgabenkontext messbar ist. Neben der Beobachtung von Bedienerereignissen wird im Folgenden ein besonderer Fokus auf die Blickbewegungsmessung gelegt, da er für die vorliegende Arbeit das zentrale Messverfahren darstellt.

### Bedienerereignisse

Ein in der Regel einfach umzusetzendes Verfahren ist die Beobachtung von Bedienerereignissen, da diese bei software-basierten Systemen ohne zusätzliches Messverfahren verfügbar sind. Ein Anwendungsfall, dem alle Menschen bei der Benutzung des Internets ausgeliefert sind, ist die Ableitung von Aufmerksamkeit basierend auf Mausposition und -klicks. Hier konnte der Zusammenhang zwischen Bedieneraufmerksamkeit und Mausinteraktion in vielen empirischen Studien nachgewiesen werden (Arapakis & Leiva, 2020; Egner et al., 2018). Die Korrelation zwischen Blick und Mausclicks ist so stark ausgeprägt, dass diese Information genutzt werden kann um Blickbewegungssysteme implizit zu kalibrieren (Jeff Huang et al., 2012; Papoutsaki et al., 2016). Die Validierung dieser bedienungs-basierten Aufmerksamkeitsabschätzung basiert wiederum oft auf Blickbewegungsmessung, welche im Folgenden diskutiert werden sollen.

### Blickbewegungsmessung

Die Blickbewegungsmessung ist das etablierteste Verfahren zur Messung von visueller Aufmerksamkeit im Umgang mit technischen Systemen. Die Verwendung dieses Verfahrens liegt die sogenannte *Eye-Mind*-Hypothese zugrunde, bei der angenommen wird, dass sich die visuelle Aufmerksamkeit mit den Informationen überschneidet, die gerade fixiert werden. Diese Annahme gründet auf Arbeiten im Zusammenhang mit Blickbewegungen beim Lesen, bei dem die Probanden bei hoher Leseaufgabenschwierigkeit länger auf einem Wort verweilen (Just & Carpenter, 1980). In weiteren Studien konnte jedoch nachgewiesen werden, dass der Mensch durchaus in der Lage ist, seine Aufmerksamkeit zu verschieben ohne den Blick dabei zu bewegen (Underwood & Everatt, 1992). Trotz dieser Limitierung gilt die *Eye-Mind*-Hypothese vor allem in Arbeitssituationen, in denen die Aufmerksamkeit durch erfahrene Bediener zielgerichtet verteilt wird, da die Position von relevanten Informationen bekannt ist (Sarter et al., 2007). Da Blickbewegungsmessung eines der zentralen Methoden der vorliegenden Arbeit ist, wird dieses Verfahren im Folgenden genauer diskutiert.

Menschliche Blickbewegungen dienen dem Zweck, den hochauflösenden Teil der Netzhaut beider Augen, der Fovea, auf ein Ziel auszurichten. Ohne genauer auf die Physiologie des Auges einzugehen (siehe hierzu Duchowski (2017)) sind dabei drei verschiedene Arten von Blickbewegungen relevant: Zunächst die **Sakkade**, eine schnelle Bewegung, um die Fovea auf ein neues Ziel auszurichten. Während der Neuausrichtung, welche je nach Distanz zwischen 10 und 100 Millisekunden dauert, ist

die visuelle Wahrnehmung stark eingeschränkt (Duchowski, 2017). Wenn der Blick auf einer Position augenscheinlich ruht, spricht man von einer **Fixation**. Schon in den Anfängen der Blickbewegungsmessung wurde nachgewiesen, dass das Auge auch bei einer Fixation nicht vollständig ruht, sondern durch Mikrosakkaden, Drift und Tremor ständig in Bewegung ist. Ein physiologischer Grund hierfür ist, dass die Wahrnehmungsfähigkeit der Netzhaut bei einem perfekt stabilisiertem Bild nach kurzer Zeit nachlässt. Dies liegt daran, dass die Rezeptoren der Netzhaut besonders empfindlich auf Lichtveränderungen reagieren (Martinez-Conde et al., 2006; Martinez-Conde & Macknik, 2015). In den meisten Anwendungen von Blickbewegungsmessungen sind diese kleinen Bewegungen des Auges jedoch von geringerer Bedeutung und werden meist unter einer Fixation zusammengefasst. Bei einer solchen Fixation, die zwischen 200 und 350 Millisekunden dauert, findet die visuelle Wahrnehmung und Verarbeitung statt, was sie zur zentralen Messgröße der Aufmerksamkeitsabschätzung macht (Feuss, 2013). Neben Sakkaden und Fixation gibt es noch die Folgebewegungen (engl. *Smooth Pursuit*) des Auges bei der Verfolgung eines sich bewegenden Blickziels.

Blickbewegungsmessungen wurden bereits in einer Vielzahl von Studien im Luftfahrtkontext eingesetzt, von denen ein Großteil von Peißl et al. (2018) und Ziv (2016) zusammengefasst sind. Obwohl es nicht immer explizit formuliert wird, werden die Mehrheit der Studien unter Annahme der *Eye-Mind*-Hypothese ausgewertet, das heißt die Verteilung der Fixationen meist als Maß für die Verteilung der visuellen Aufmerksamkeit gesehen. Beispielsweise argumentiert Wickens (2015), dass die Arbeitsplätze in der Luftfahrt so stark durch visuelle Aufgaben dominiert werden, dass die Blickposition ein gutes Ersatzmaß für die visuelle Aufmerksamkeit des Operateurs ist. Zentral für den Einsatz von Blickbewegungsmessung ist die Definition von Interessensgebieten (engl. *Areas of Interest*, Aoi), die die für den Bediener relevanten Informationsquellen und Bedienoberflächen räumlich beschreiben. Die verschiedenen Blickbewegungsmetriken werden dann den Interessensgebieten zugeordnet, wie beispielsweise die Fixationsfrequenz auf dem Interessensgebiet Kartendisplay (Wickens, 2021). Abbildung 2-4 zeigt eine beispielhafte Folge von Fixationen zwischen vordefinierten Interessensgebieten im Cockpit einer Boeing 747.

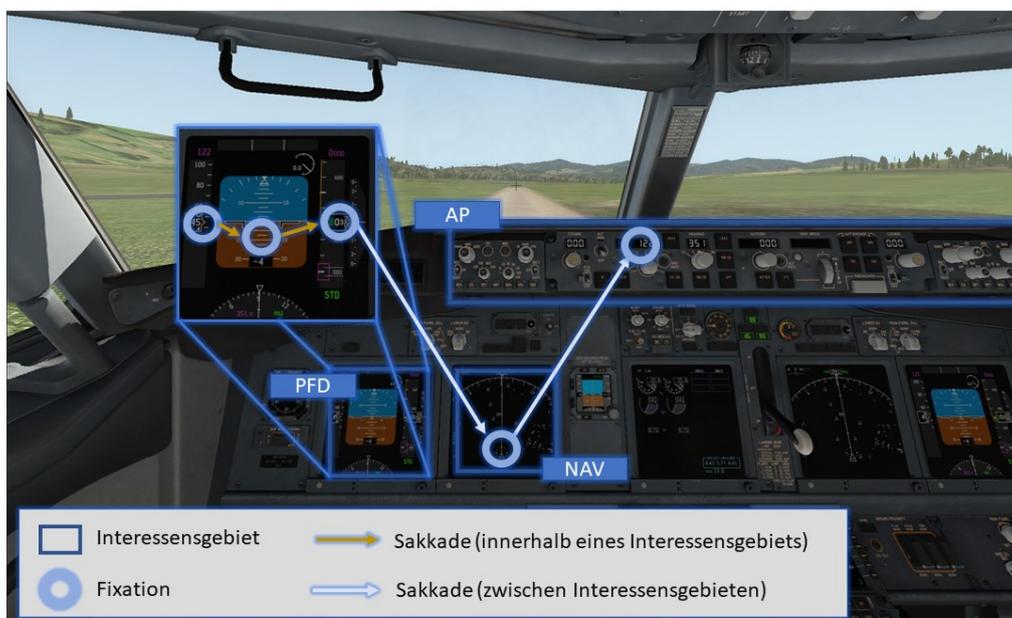


Abbildung 2-4 - Schematische Abbildung von Fixationsfolgen innerhalb und zwischen Interessensgebieten, welche durch Sakkaden verbunden sind.

Aufbauend auf Fixation und Sakkade werden im Großteil der Studien folgende Metriken verwendet (Glaholt, 2014):

- **Fixationsdauer, -anzahl und -frequenz:** Eine Fixation wird durch Blickbewegungssysteme durch eine niedrige Winkelgeschwindigkeit der gemessenen Blickachse klassifiziert. Die Fixationsanzahl bezeichnet die Zahl der Fixation auf ein Interessensgebiet und wird meist als relative Größe in Bezug auf die Gesamtanzahl der Fixationen angegeben. Die Fixationsfrequenz ist definiert als Fixationszahl auf ein Interessensgebiet pro Sekunde.
- **Sakkadenanzahl und -länge:** Sakkaden sind definiert durch die abrupte Bewegung des Blickes mit hohen Winkelgeschwindigkeiten. Die Sakkadenlänge ist die zurückgelegte Distanz von Start- zu Endpunkt.
- **Verhältnis von Sakkaden- zu Fixationsdauern:** Wenn Sakkaden für die Suche nach Informationen und Fixation für die Verarbeitung stehen, dann kann das Verhältnis ein Indikator für den Fokus den Bedieners sein (Goldberg & Kotval, 1999). Dehais et al. (2015) erweiterten diese Idee mit der *Explore-Exploit-Ratio*, welche das Zeitverhältnis kurzer Fixationen und Sakkaden zu langen Fixationen abbildet. Damit wurden in einem Flugsimulatorexperiment Situationen identifiziert, in denen Versuchsteilnehmer von einer Modusänderung des Autopiloten überrascht wurden und nach einem Fehler suchten.
- **Verweilzeit & proportionale Verweilzeit:** Die Verweilzeit (engl. *dwel time*) ist die Dauer aller aufeinanderfolgenden Fixationen auf ein Interessensgebiet. Wenn Interessensgebiete viele Informationen umfassen, wie beispielsweise eine Kartenanzeige, werden meist mehrere Orte im Interessensgebiet fixiert. Die summierte Dauer dieser Fixationen ist die Verweilzeit und die proportionale Verweilzeit wird normiert durch die gesamte Verweilzeit. Sie wird oft verwendet, um die grundsätzliche Verteilung von visueller Aufmerksamkeit zu beschreiben.

Eine große Anzahl von Studien beschäftigt sich mit der grundlegenden Verteilung von Blickbewegungen im Cockpit. Dabei allokierten Piloten ziviler Flugzeuge grundsätzlich mehr visuelle Aufmerksamkeit auf Cockpitinstrumente (ca. 60-70%) als auf die Außenwelt (ca. 30-40%) (Colvin et al., 2005; Helleberg & Wickens, 2003). Sarter et al. (2007) zeigten in einer Simulatorstudie mit 20 Piloten, dass das Primary Flight Display (31%) und die Kartenansicht (25%) den größten Anteil der Blickverweilzeiten haben, wobei die Verteilung grundsätzlich von der Flugphase und damit von der Aufgabensituation abhängt. Beim Vergleich der Blickbewegung von erfahrenen Piloten mit Novizen in der zivilen Luftfahrt konnte mehrfach gezeigt werden, dass eine hohe Erfahrung mit niedrigerer Blickverweilzeit aber höherer Fixationsfrequenz auf aufgabenrelevanten Interessensgebieten einhergeht (Brams et al., 2018; Lounis et al., 2021).

Während die meisten der Arbeiten die foveale Wahrnehmung, also die exakte Position der Blickfixationen, auswerten, beschäftigten sich drei Studien mit der peripheren Sicht. Die Netzhaut außerhalb der Fovea ist nicht blind, sondern mit einer abnehmenden Zahl von Zapfen und einer höheren Zahl von Stäbchenrezeptoren ausgestattet, die besonders lichtempfindlich sind. Das Blickfeld des menschlichen Auges umfasst horizontal annähernd 180° und vertikal etwa 130°. Der Bereich der höchsten Sehschärfe ist ein exzentrischer Bereich von 2° um die Fovea. Dieser Bereich ist in hoher Dichte mit dem farbsensitiven Rezeptor, den sogenannten Zapfen, ausgestattet. Dieser Bereich ist

besonders gut geeignet für die Wahrnehmung von Bewegung im peripheren Sichtfeld (Duchowski, 2017, S. 30; Strasburger et al., 2011). Dass das periphere Sehen auch für Wahrnehmung von Informationen im Flugzeugcockpit relevant ist, konnte in einer Studie von J. Kim et al. (2010) nachgewiesen werden. Dabei blickten Piloten in einer simulierten Landung entgegen der Annahme der Autoren selten auf einen virtuellen Indikator, der den Landepunkt des Flugzeugs vorhersagte, obwohl sie das Flugzeug genauso kontrollierten, dass dieser Indikator auf dem Landepunkt lag. Die Autoren schlossen daraus, dass der Indikator im peripheren Sichtfeld wahrgenommen wurde. Ähnliche Ergebnisse erzielten Schaudt et al. (2002) beim Vergleich einer klassischen Gestaltung eines Fahrtmessers mit einer Variante, die die Wahrnehmung im peripheren Sichtbereich ermöglichen sollte. Die Abbildung 2-5 zeigt die konventionelle Variante, die eine Fixation notwendig macht gegenüber der „peripheren Gestaltung“, die die Richtung und den Betrag der Abweichung von der gewünschten Geschwindigkeit durch Größe und Orientierung der Pfeile darstellt. Die Autoren konnten feststellen, dass die Information im Peripheren wahrgenommen werden konnte und damit auch eine Leistungssteigerung in einem Experimentalszenario einherging. In einem Versuch mit 20 Studenten und 20 Fluglehrern wurde mithilfe einer Blickbewegungsmessung alle Cockpitinformation in der Peripherie der Blickposition entfernt. Dies hatte zur Folge, dass die Leistung der Versuchsteilnehmer sank, wobei die Fluglehrer ein größeres Problem mit dem Fehlen der Informationen im peripheren Sichtfeld hatten als die Novizen (Fox et al., 1996).

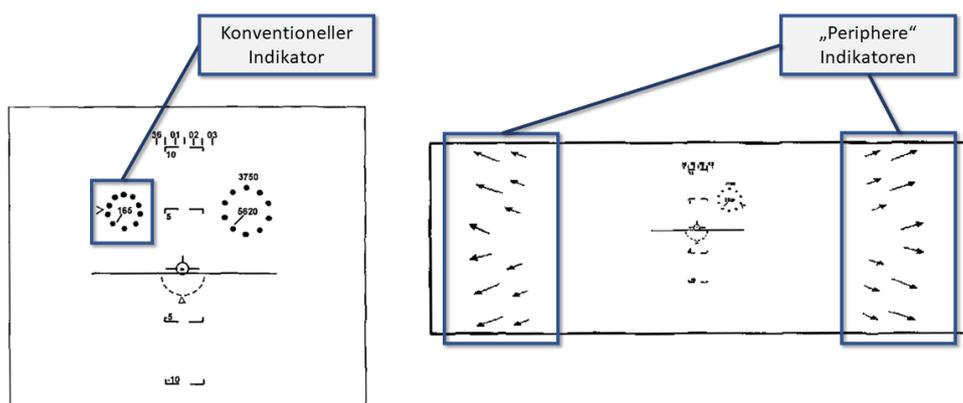


Abbildung 2-5 - Vergleich der Gestaltung zweier unterschiedlicher Fahrtmesser in einem HUD (Schaudt et al., 2002): Konventionell (links) gegenüber der neuartigen Gestaltung (rechts)

Um das dynamische Blickbewegungsverhalten in Arbeitssituationen zu erfassen, gibt es Metriken, die Transitionen von einem zum nächsten Interessensgebiet und auftretende Reihenfolgen in der Überwachung abbilden (Lounis et al., 2021):

- Transitionsmatrix und Transitionsmatrixdichte:** Transitionsmatrizen können ausgehend von einem Interessensgebiet die Wahrscheinlichkeit angeben, welche Gebiete als nächstes fixiert werden. (Wickens, 2015) erstellte mit seinem SEEV-Modell Transitionsmatrizen, um Blickpfade von Piloten zu modellieren und vorherzusagen. Die Transitionsmatrixdichte kann aus dieser Darstellung errechnet werden, indem die Anzahl der Einträge der Matrix ungleich 0 durch die Gesamtzahl der Einträge geteilt wird. Eine niedrige Dichte spricht beispielsweise für ein Überwachen von einzelnen Interessensgebieten. Je nach Situation kann dies als fokussiert oder als vernachlässigend gegenüber anderen Interessensgebieten gewertet werden (Goldberg & Kotval, 1999).

- **K-Koeffizient:** Ein weiteres Berechnungsverfahren das anhand Blickbewegungen zwischen fokussierter und verteilter Aufmerksamkeit zu unterscheiden versucht, ist der K-Koeffizient, welcher ein Verhältnis von normierter Fixationsdauer zu normierter Sakkadenlänge beschreibt. Dieser wurde beispielsweise genutzt, um zwischen paralleler und serieller Suche zu unterscheiden (Krejtz et al., 2016). In einer modifizierten Form mit Verweilzeiten statt Fixationsdauern wurde der K-Koeffizient verwendet, um zwischen fokussierter und verteilter Aufmerksamkeit eines Piloten in verschiedenen Flugphasen zu unterscheiden (Lounis, Hassoumi et al., 2020).
- **Blickbewegungsentropie & -komplexität:** Aufbauend auf Shannon's Maß für Informationsgehalt, quantifiziert die Blickbewegungsentropie die Vorhersagbarkeit der Blickbewegung (Shiferaw et al., 2019). Eine hohe Entropie steht für ein zufälliges, hochfrequentes Wechseln von Interessensgebieten, während eine niedrige Entropie eine vorhersagbare Überwachungsstrategie darstellt. Ein damit verwandtes Maß ist die Lempel-Ziv Komplexität, die als Datenkompressionsalgorithmus entwickelt wurde und wurde für Blickbewegungsmessung im Simulatorcockpit angewendet (Lounis, Peysakhovich & Causse, 2020b).

### 2.1.5 RESÜMEE

Die zu Beginn des Kapitels diskutierten Arbeiten zeigen, dass der Begriff Aufmerksamkeit verschiedene Aspekte menschlicher Informationsverarbeitung unter sich vereint. In den Modellen wird Aufmerksamkeit entweder als Selektionsmechanismus oder als Ressource verstanden, die zur Informationsverarbeitung gebraucht wird. Im Verlauf dieses Kapitels wurde der Selektionsmechanismus genauer beleuchtet, der relevante Informationen zielorientiert zur Verarbeitung auswählt und in verschiedenen Aufgabentypen eine zentrale Rolle spielt. Der darauffolgende Abschnitt beschrieb unterschiedliche Verfahren zur Messung von Aufmerksamkeit. Grundsätzlich kann hier zwischen Methoden unterschieden werden, die Leistung, physiologische Messgrößen oder die Bedienung des Systems messen. Nun ist die Frage, welche dieser Methoden sich am besten zur Verwendung in einem online-fähigen Messsystem von Aufmerksamkeit eignen.

Die Übertragung leistungsbasierter Messmethoden auf eine komplexe Arbeitsumgebung wie beispielweise ein Flugzeugcockpit gestaltet sich schwierig. Zwar sind die beschriebenen Leistungsmaße in einer Laborumgebung einfach zu messen, in naturalistischen Arbeitssituationen ist dagegen nicht immer direkt klar, was das Leistungsmaß sein soll, das für die Aufmerksamkeit des Bedieners steht. Zusätzlich besteht die Gefahr, dass Leistungsdekremente erst zeitversetzt messbar sind. Sauer et al. (2017) argumentieren aus diesem Grund auch, dass leistungsadaptive Systeme in realistischen Arbeitsumgebung nicht gut funktionieren. So kann der Pilot eines zivilen Flugzeugs beispielweise mehrere Ziele gleichzeitig verfolgen, wie das Umfliegen eines Gewitters bei gleichzeitiger Minimierung der Turbulenzen und Beschleunigung für die Fluggäste und kontinuierlicher Kommunikation mit der Luftverkehrskontrolle. Die Verarbeitung physiologischer Signale eignet sich zur zeitnahen Bestimmung von kognitiven Zuständen wie z.B. Müdigkeit, die Einfluss auf die Aufmerksamkeit haben. Die Ergebnisse dieser Messung sind jedoch immer kontextfrei und können keine Information darüber geben, welcher Information gerade Aufmerksamkeit geschenkt wird. Die Bedienungsbeobachtung kann hier mehr Kontextinformationen liefern, ist aber stets abhängig vom technischen System. Im Zusammenhang mit visueller Aufmerksamkeit ist die Blickbewegungsmessung ein etabliertes Messverfahren, deren Metriken im Zusammenhang mit

Interessensgebieten des technischen Systems beziehungsweise im Aufgabenkontext des Bedieners formuliert werden können.

Im Abschnitt über die Blickbewegungsmessung wurden die gängigen Metriken zur Auswertung der Blickbewegungsmessung, wie beispielsweise Verweilzeit oder Fixationsdauer, und darauf aufbauende Größen wie Transitionsmatrizen diskutiert. Dabei wurden Studien vorgestellt, die Blickbewegungsmessung im Flugzeugcockpit eingesetzt haben, um die Verteilung der Aufmerksamkeit zu analysieren oder Unterschiede zwischen Novizen und Experten zu erkennen. Die dort vorgestellten Metriken besitzen eine akkumulierende Natur, welche Einblick in das Makro-Level des Pilotenverhaltens ermöglichen und sich für die Identifikation von grundsätzlichen Korrelationen zwischen Blickverhalten und Flugleistung zu identifizieren. Ein Problem bei der Verwendung dieser Metriken in einem adaptiven System ist, dass sie keine situative Aussagekraft besitzen. Sie können nichts darüber aussagen, welche Informationen zu gegebenem Zeitpunkt verarbeitet worden sind.<sup>2</sup>

Um diese situative Aussagekraft zu gewährleisten, muss Aufmerksamkeitsabschätzung im Flugzeugcockpit als Klassifikationsproblem mit einer zentralen Frage gesehen werden: Welchen Informationen schenkt der Pilot zum vorliegenden Zeitpunkt seine Aufmerksamkeit? Bevor hierfür ein Konzept vorgestellt wird, wird zunächst das Modell von Situationsbewusstsein betrachtet, da es eine Brücke zwischen dem kognitiven Konstrukt der Aufmerksamkeit und der situativen Aufgabenumgebung eines Piloten schlagen kann.

## 2.2 SITUATIONSBEWUSSTSEIN

Seit einigen Jahren wird der Begriff **Situationsbewusstsein** (SB, engl. *Situation Awareness*) zunehmend von Vertretern der Luftfahrtindustrie (vergleiche F-35 Website (2020)), Piloten und Wissenschaftlern verwendet (Stanton et al., 2017). Dabei steht der Begriff für das Situationsbild eines Bedieners in seinem hoch-dynamischen Arbeitsumfeld, umgangssprachlich auch „*big picture*“ genannt. SB ist hochrelevant für die Leistung eines Bedieners, da dieser mit einem umfassenden Situationsbild in der Lage ist, seine Aufmerksamkeit auf die wichtigste Aufgabe zu lenken, Informationen im System einzuordnen und die Dynamik einer Situation zu antizipieren. In hochautomatisierten Systemen gewinnt das Thema SB an Bedeutung, da der Bediener nicht mehr in der Ausführung jeder Aufgabe involviert ist und damit automatisch ein hohes SB hat, sondern sich dieses durch Überwachung des Systemzustands und Kenntnis der automatisierten Funktionen selbst „erarbeiten“ muss.

Trotz des steigenden Interesses an SB, gibt es noch keine unstrittige Definition und damit auch keine verlässliche Methode zur Messung von SB. Ein Grund hierfür ist, dass die intuitive Vorstellung eines „*big picture*“ viele Aspekte aus der Kognitionspsychologie und der Arbeitswissenschaft umfasst. Hierzu zählt die menschliche Informationsverarbeitung aber auch Domänenwissen über die relevanten Ziele und Aufgaben einer Arbeitssituation. Dass die Modellvorstellung von Situationsbewusstsein all diese Aspekte umfasst, führt dazu, dass der Begriff in vielen Kontexten ohne strenge Definition verwendet wird. Die folgende Diskussion gibt einen Überblick über verschiedene SB-Modelle und stellt im Anschluss verschiedene Methoden zur Messung von SB vor.

Parallel zu den Modellen von Situationsbewusstsein des individuellen Bedieners, gibt es Arbeiten über das Situationsbewusstsein von Teams (Salas, 2017) oder gesamten sozio-technischen Systemen (Stanton, 2016). Diese Überlegungen haben in der Forschungsliteratur vorwiegend einen theoretischen und analytischen Charakter und bieten noch keine etablierten Messverfahren. Da der

---

<sup>2</sup> Eine Ausnahme bilden die Arbeiten mit EEG zur Klassifikation von nicht verarbeiteten auditiven Signalen Dehais et al. (2019); Klaproth et al. (2020). Den Autoren ist aber keine Arbeit für visuelle Aufmerksamkeit bekannt.

Fokus der Arbeit auf einem Assistenzsystem für einen individuellen Bediener liegt, werden im Folgenden nur Modelle von individuellem SB diskutiert.

### 2.2.1 MODELLE

Eine Schwierigkeit bei der Definition von Situationsbewusstsein liegt in der Gefahr, entweder eine zu allgemeine Definition zu wählen, welche keine Aussagen treffen kann, oder eine zu situationsabhängige Definition zu beschreiben, sodass keine allgemeinen Aussagen möglich sind. Es ist beispielsweise offensichtlich, dass der Mensch *etwas* wissen muss, um seine Aufgabe zu erfüllen. Ebenso wenig nützlich ist es, Situationsbewusstsein im Bezug zu einer Situation unter speziellen Bedingungen, speziellen Ereignissen und spezieller Dynamik zu definieren (Sarter & Woods, 1991). Es handelt sich bei Situationsbewusstsein um ein *Framework*-Modell, welches Erkenntnisse aus der Kognitionspsychologie und der Ergonomie mit Wissen über die Aufgaben einer Arbeitssituation zusammenführt. Das macht es schwierig, das Modell als Ganzes zu beweisen, da es nicht direkt falsifizierbar ist (Endsley, 2015).

Als Ausgangspunkt kann die Überlegung dienen, welche Aspekte ein Modell von SB haben muss, um zur Erklärung menschlichen Verhaltens und Fehler bei der Bedienung von technischen Systemen beitragen zu können. Hierfür muss das Modell den **Menschen als Bediener und dessen kognitive Prozesse** der Entscheidungsfindung und Durchführung beinhalten. Ein weiterer Bestandteil muss **die Arbeitsdomäne** sein, da hier definiert wird, welche Bestandteile der Umgebung überhaupt relevant für das Ziel des Bedieners sind. Um in einem technischen System Anwendung zu finden, sollte ein Modell zuletzt nicht rein auf theoretischer Überlegung basieren, sondern Möglichkeiten zur **Operationalisierung** von Messgrößen ermöglichen. Unter Berücksichtigung dieser Anforderungen werden im Folgenden drei Kategorien von Situationsbewusstseinsmodellen beschrieben. Ziel der Literaturrecherche ist ein Überblick über die theoretischen und experimentellen Arbeiten in diesem Kontext. Außerdem werden die gemeinsamen Aspekte verschiedener Modelle herausgearbeitet. Ein besonderes Augenmerk liegt auf den beobachtbaren Elementen von Situationsbewusstsein, da diese der vorliegenden Arbeit eine wichtige Rolle spielen.

#### 2.2.1.1 Informationsverarbeitungsparadigma

Das Modell für Situationsbewusstsein von Endsley ist das populärste Modell in der Forschungsliteratur (Endsley, 1995b). Im Zentrum dieses Modells stehen drei Ebenen (engl. *Level*), mithilfe derer Situationsbewusstsein definiert wird als

“THE PERCEPTION OF THE ELEMENTS IN THE ENVIRONMENT WITHIN A VOLUME OF TIME AND SPACE, THE COMPREHENSION OF THEIR MEANING, AND THE PROJECTION OF THEIR STATUS IN THE NEAR FUTURE” (ENDSLEY, 1995B).

Die drei Ebenen Wahrnehmung, Verständnis und Projektion beschreiben die zentralen Elemente des Modells und sind wie folgt definiert:

#### **EBENE 1: WAHRNEHMUNG**

Die erste Voraussetzung für ein gutes Situationsbewusstsein ist, dass ein Bediener alle relevanten Informationen in dem von ihm kontrollierten System wahrnimmt. Mit relevanten Informationen sind Systemzustände, Eigenschaften und Dynamik von Elementen der Umgebung gemeint, deren Relevanz durch das Ziel der Arbeitsdomäne festgelegt wird. Für Piloten ziviler Luftoperationen ist zum Beispiel die aktuelle Position des Flugzeugs relativ zum Boden und zugewiesenen Luftraum relevant. Die Wahrnehmung ist also der Informationsgewinn durch Reize der Umgebung, wie beispielsweise das

Lesen einer Systeminformationen auf digitalen Anzeigen oder das Spüren von Turbulenzen durch die Vibrationen des Flugzeugs.

### EBENE 2: VERSTÄNDNIS

Die zweite Ebene beschreibt das Verständnis der wahrgenommenen Informationen. Verständnis steht hier zum einen für die Relation der wahrgenommenen Informationen zu den Zielen und Entscheidungen, die ein Bediener treffen muss, sowie zum anderen für die Kombination von Informationen, um Schlüsse zu ziehen. Dabei spielen *mentale Modelle* eine zentrale Rolle, mit deren Hilfe ein Bediener in der Lage ist, Informationen in das bestehende Situationsbild zu integrieren, normale von abnormalen Zuständen zu unterscheiden und zukünftige Zustände zu antizipieren.

### EBENE 3: PROJEKTION

Der dritte Bestandteil des Situationsbewusstseins ist die Projektion zukünftiger Systemzustände. Diese Art des „Voraussehens“ erfordert Erfahrung sowie ein gutes mentales Modell der Situationsdynamik. Mittels einer Projektion ist ein Bediener auf dynamische Änderungen vorbereitet und kann Entscheidungen sowie Aktionen projizieren. So kann ein erfahrener Autofahrer die Zeit einer Ampelperiode grob abschätzen und sich beispielsweise schon auf das Bremsen beziehungsweise Beschleunigen vorbereiten. Laut Endsley benötigt der Aufbau der Projektionsebene große mentale Anstrengungen und findet erst statt, wenn den niedrigeren Ebenen genüge getan wurde.

Situationsbewusstsein ist also die Grundlage, auf der Entscheidungen bei der Bedienung eines dynamischen Systems getroffen werden. Wie Abbildung 2-6 verdeutlicht, sind dabei verschiedene kognitiven Prozesse beteiligt wie Informationsverarbeitung, Langzeitgedächtnis, Automatismus aber auch die Fähigkeiten und Erfahrungen, die diese kognitiven Prozesse optimieren. Darüber hinaus sind Systemfaktoren wie die Fähigkeiten des Systems, die Mensch-Maschine-Schnittstelle, Komplexität und Automation als Einflussfaktoren definiert.

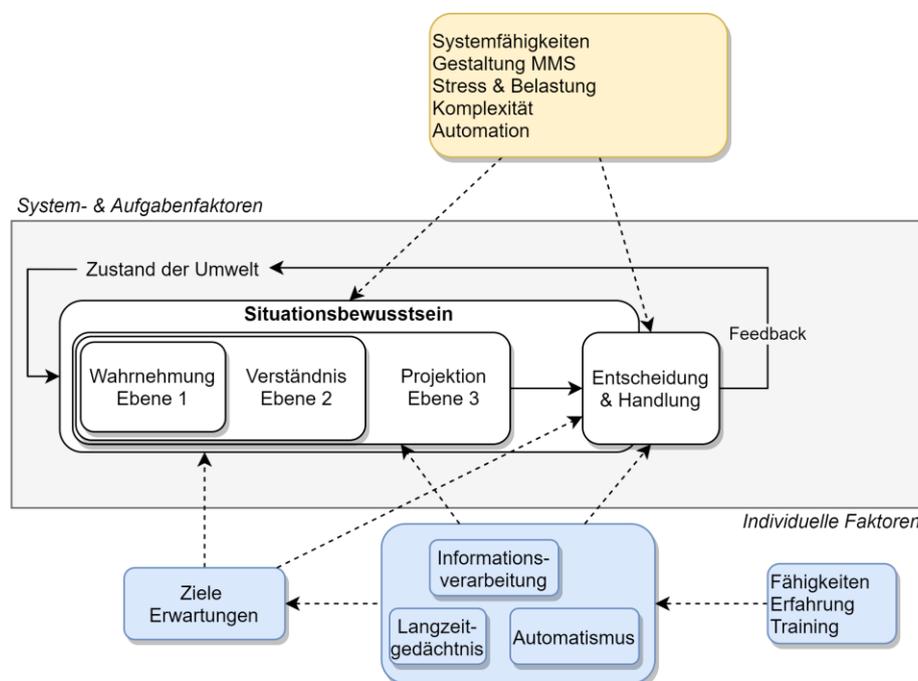


Abbildung 2-6 - Prozess und Einflussfaktoren bei der Entscheidungsfindung eines Bedieners (Endsley & Jones, 2012)

Das Resultat des Prozesses ist eine Entscheidung des Bedieners und eine Ausführung der notwendigen Aktionen. Ein besseres Situationsbild führt laut diesem Modell auch zu einer besseren und schnelleren Entscheidung. Die Ausführung kann fehlerbehaftet sein (vergleiche *Slips* in Abschnitt 2.3.1) obwohl das Situationsbewusstsein des Bedieners gut ist. Aus ihrem Modell leiten die Autoren verschiedene Richtlinien zur Gestaltung von Mensch-Maschine-Schnittstellen ab, welche die Bildung von Situationsbewusstsein unterstützen (Endsley & Jones, 2012). Teil dieser Leitfäden ist eine an Zielen orientierte Aufgabenanalyse (engl. *Goal Directed Task Analysis*, GDTA), deren Ziel es durch Befragung von erfahrenen Operateuren ist, die kritischen Informationen einer Arbeitsdomäne zu identifizieren, die für die Entscheidungsfindung wichtig sind. Ferner bietet dieses Modell Methoden zur Messung von Situationsbewusstsein, die im weiteren Verlauf dieses Kapitels noch diskutiert werden.

### 2.2.1.2 Ökologisches Paradigma

Das Ebenenmodell von Endsley ist geprägt durch das kognitive Modell der Informationsverarbeitung (siehe 2.1.1) und stellt Situationsbewusstsein als linearen Verarbeitungsprozess inklusive Ein- und Ausgangsgrößen dar. Ein Gegenentwurf hierzu liefern Theorien, die auf der Schematheorie von (Neisser, 1976) fußen, welches Wahrnehmung und Aktion eines Bedieners mit einem zyklischen Prozess erklärt, der auf sogenannten *Schemata* basiert und in Abbildung 2-7 dargestellt ist. Ein Schema bezeichnet den gedanklichen Rahmen, in dem neue Informationen und Ereignisse interpretiert werden. Der zyklische Prozess ist die explorative Wahrnehmung des Bedieners, die durch das Schema gesteuert wird und welche es wiederum mit neuen Informationen und Interpretationen aktualisiert oder verändert. Das Modell von Adams et al. (1995) basiert auf dieser Vorstellung und betont die zyklische Natur von Situationsbewusstsein. In den Augen der Autoren kann Situationsbewusstsein sowohl als Zustand als auch als Prozess gesehen werden, da das Bewusstsein über eine Situation auch den Prozess der Wahrnehmung und damit auch die Ausrichtung der Aufmerksamkeit beeinflusst.

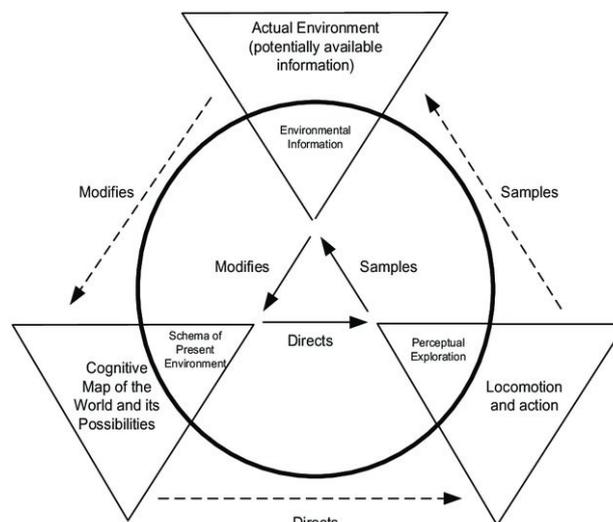


Abbildung 2-7 - Der Wahrnehmungszyklus nach Adams et al. (1995)

Ähnlich wie Adams, definieren Kip und Hancock (1995) Situationsbewusstsein als das nach außen gerichtete Bewusstsein, das der Adaption des Verhaltens zur Erfüllung extern vorgegebener Ziele dient. Ein gutes Situationsbewusstsein leitet den Bediener bei der Verteilung seiner Aufmerksamkeit. Das Verhalten, das durch gutes Situationsbewusstsein generiert wird, ist

“THE AGENT'S SOLUTION TO THE PROBLEM OF KNOWING THOSE CUES AND DEMANDS IN THE ENVIRONMENT THAT ENABLE IT TO TAKE ACTION THAT ALIGNS WITH THE DICTA OF THE ARBITER OF PERFORMANCE.” (KIP & HANCOCK, 1995)

Dabei wird betont, dass eine externe Definition von Zielen Voraussetzung dafür ist, dass man Situationsbewusstsein sinnvoll operationalisieren kann. Dies steht im Kontrast zum Drei-Ebenen-Modell, dass die verfolgten Ziele als interne Größe des menschlichen Bedieners modelliert. Eine weitere Kritik der Autoren ist, dass Endsley's Modell sich zu sehr auf den Bediener konzentriert und zu wenig auf seine Interaktion mit dem Arbeitsplatz. Ein interessanter Aspekt dieses Modells ist jedoch die Definition einer „unveränderlichen Größe“ (engl. *Invariant*) im Kern des verhaltensgenerierenden Situationsbewusstseins. Diese Größe soll das Verhalten eines Bedieners in einer dynamischen Umgebung operationalisieren. Dies wird mittels eines „Risiko-Raums“ (engl. *Risk-Space*) festgelegt. Abbildung 2-8 zeigt ein Beispiel aus der Domäne der Fluglotsen. Die Achsen bilden die relevanten Informationen für die Entscheidung, ob die horizontale Luftraumtrennung zweier Flugzeuge ausreichend ist oder der Konflikt durch den Fluglotsen gelöst werden muss. Diese Operationalisierung kann herangezogen werden, um die Entscheidung eines Fluglotsen vorherzusagen und zu definieren, worauf der Fluglotse bei der Durchführung der Aufgabe achten sollte.

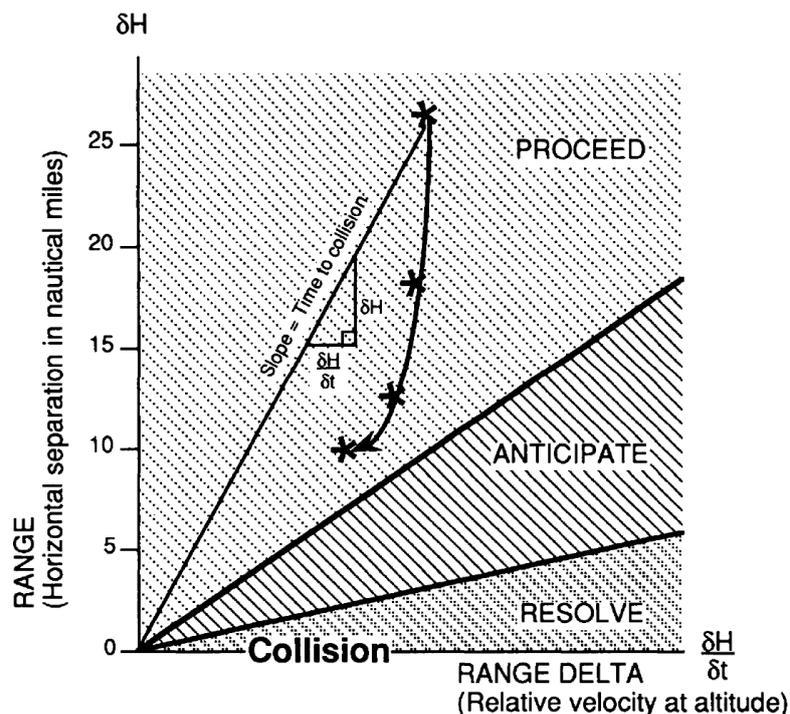


Abbildung 2-8 - Risikoraum eines Fluglotsen bei der Überwachung der räumlichen Trennung zweier Flugzeuge (Kip & Hancock, 1995)

Ein alternatives Modell verwendet nicht den Begriff Situationsbewusstsein, sondern spricht vom *Sensemaking* (Klein, 1993; Klein et al., 2007). Im Zentrum dieser Theorie stehen dabei mentale Denkraum (engl. *Frames*), die verwendet werden, um die Informationen der Umgebung zu interpretieren, zu filtern und ihnen eine Bedeutung zuzuweisen und Beziehungen zu definieren. Ausgehend von einem initialen Modell, suchen Bediener immer das beste Modell zur Erklärung der auftretenden Daten. Ähnlich wie bei den anderen Modellen wird die zyklische Natur des Prozesses betont, in der die mentalen Modelle die Aufmerksamkeit und Informationsverarbeitung des Bedieners steuern. Denkraum können Situationen organisieren in Form von Karten, Geschichten (kausale

sowie zeitliche Zusammenhänge und Situationen) oder funktionale Abläufe. Im Gegensatz zu den anderen Modellen gibt es hier aber keine Anhaltspunkte für eine messtechnische Operationalisierung dieses Modells.

### 2.2.1.3 Rechnermodelle

Basierend auf den theoretischen Arbeiten werden in der Forschungsliteratur auch einige Rechnermodelle beschrieben, welche zum Ziel hatten, spezifische Aspekte des Situationsbewusstseins in einem Softwaremodell abzubilden. Die im Folgenden beschriebenen Arbeiten basieren alle auf der informationsverarbeitenden Sichtweise. Ein Grund hierfür könnte sein, dass das zugrundeliegende Modell der menschlichen Informationsverarbeitung am ehesten der rechnergestützten Datenverarbeitung entspricht.

Wickens et al. (2008) entwickelten das Rechnermodell *Attention and Situation Awareness (A-SA)*, das die Wahrscheinlichkeit von Fehlern aufgrund fehlenden Situationsbewusstseins vorhersagen sollte. Hierfür werden folgende drei kognitive Mechanismen modelliert: die Verteilung der Aufmerksamkeit, die Limitierung der Aufmerksamkeitsressourcen und die Arbeitsbeanspruchung. Das Aufmerksamkeitsmodul beschreibt, wie viel Aufmerksamkeit den auftretenden situativen Ereignissen (E) zugewandt wird. Zur Modellierung dieser Verteilung ziehen die Autoren das SEEV-Modell (siehe Abschnitt 2.1.3) heran. Gleichzeitig erzeugt die Zuwendung von Aufmerksamkeit zusätzliche Arbeitsbeanspruchung, welche die Ressourcen für Aufmerksamkeit mindern. Diese Beanspruchung nimmt zeitlich abhängig von der Anzahl auftretender Ereignisse sowie einer Zeitkonstante ab. Sobald Aufmerksamkeit auf Informationen gerichtet wird, wird das Modul zur Abschätzung von Situationsbewusstsein um einen gewichteten Wert des Ereignisses aktualisiert, wobei die Gewichtung durch den Anteil der Aufmerksamkeit auf dieses Ereignis festgelegt wird (*Belief Module*). Der Wert von Situationsbewusstsein nimmt langsamer beziehungsweise schneller ab, je nachdem ob neue Ereignisse auftreten. Der Wert des Situationsbewusstseins wirkt sich wiederum auf die Effektivität der Aufmerksamkeitsverteilung aus, sowie auf die Wahrscheinlichkeit, dass der Bediener bei einer anstehenden Entscheidung die richtige Aktion auswählt.

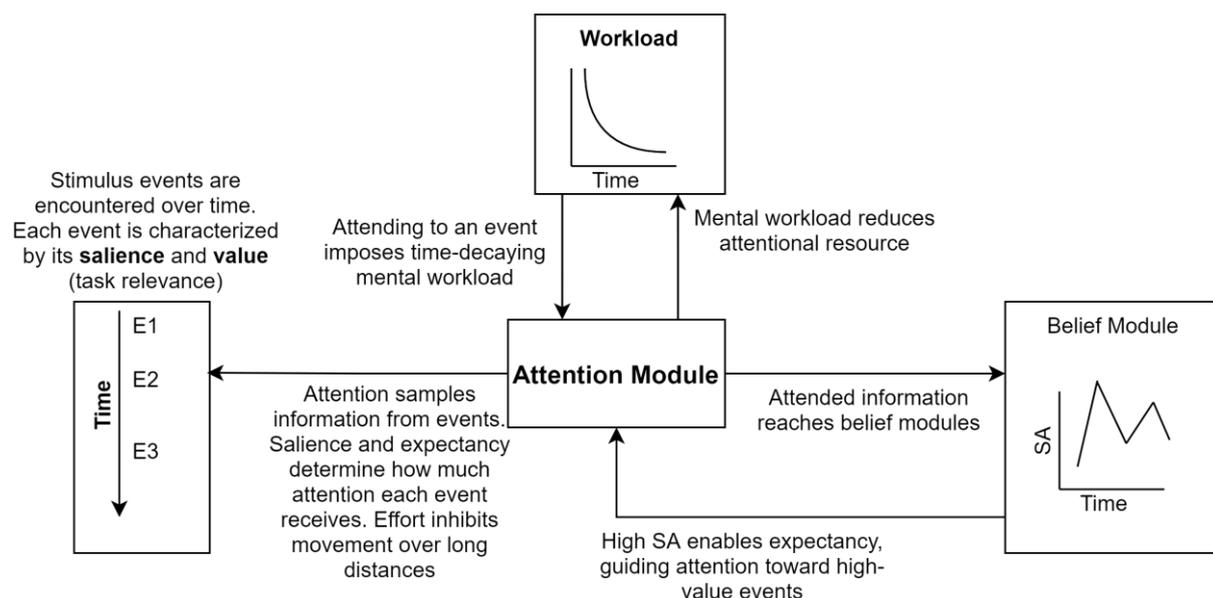


Abbildung 2-9 - Attention-SA Modell nach Wickens et al. (2008)

Die Autoren stellen drei Anwendungsfälle vor. Der erste Anwendungsfall, *Taxiing*-Szenarien in der zivilen Luftfahrt, dient der analytischen Untersuchung unterschiedlicher Zeitkonstanten. In den beiden

anderen Anwendungsfällen werden neuartige Displays untersucht, mit deren Hilfe Situationsbewusstsein verbessert werden soll. Leider können die Autoren das vorgestellte Modell nicht auf diese Anwendungen übertragen und verbleiben bei einer Validierung des Aufmerksamkeitsmoduls.

In einer anderen Studie nutzen R. E. T. Jones et al. (2011) sogenannte *Fuzzy-Cognitive-Maps* (FCM) um den Zusammenhang zwischen verschiedenen Informationselementen einer Aufgabenumgebung zu modellieren. Dabei wird eine *Goal Directed Task Analysis* (GDTA) herangezogen, um die Bestandteile der FCM zu identifizieren. Dann werden mittels Gewichtungen zwischen den einzelnen Knoten der FCM positive und negative Einflüsse von Informationen auf höher aggregierte Informationen beschrieben. Diese Gewichtungen werden dann in einem rechnergestützten Inferenzschritt genutzt, um vom initialen Zustand der Eingangsknoten auf den Wert des Situationsbewusstseins in höherwertigen Knoten zu schließen. Die Abbildung 2-10 zeigt ein konzeptionelles Bild dieser FCM. Die Autoren der Studie nutzten das Modell, um Situationsbewusstsein von Infanteriezugführern bei der Entscheidungsfindung zu beschreiben. Eine Studie, die im Kontext dieser Arbeit angefertigt wurde, griff diese Modellierung zur online-fähigen Abschätzung von Situationsbewusstsein auf (Schwerd & Schulte, 2021b).

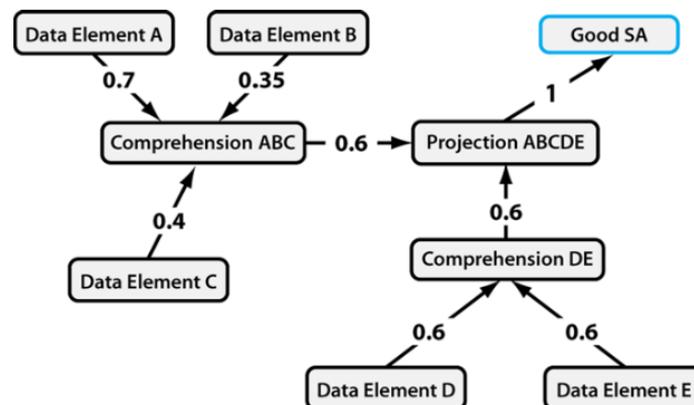


Abbildung 2-10 - Konzeptionelle *FCM Submap* aus der Arbeit von R. E. T. Jones et al. (2011)

Johnson et al. (2017) stellen eine Studie vor, in denen sinkendes Situationsbewusstsein als die wachsende Unsicherheit in Systemzuständen zwischen Blickfixationen der assoziierten Anzeigen modelliert ist. Verlauf dieser Unsicherheit ist in Abbildung 2-11 dargestellt. Die dargestellte Transition von einem zum anderen Modus der Automation lässt in diesem Beispiel die Priorität des Zustands sinken, weshalb eine höhere Unsicherheit akzeptiert wird. Diese Unsicherheit wird von den Autoren als proportional zur Genauigkeit des internen mentalen Modells des Bedieners vom Systemverhalten definiert. Mithilfe der Modellierung von Aufmerksamkeit mittels SEEV entwickeln die Autoren ein Plattform-Bediener-Modell für den Anwendungsfall eines Mondfahrzeugs, das von einem Bediener kontrolliert werden. Damit waren die Autoren in der Lage, die Verteilung von Aufmerksamkeit in Experimenten vorherzusagen und auch die Auswirkungen von höherer Arbeitsbelastung auf das Situationsbewusstsein zu zeigen. Hohe Arbeitsbelastung in einer primären Aufgabe verursachte ein sinkendes SB in Sekundäraufgaben.

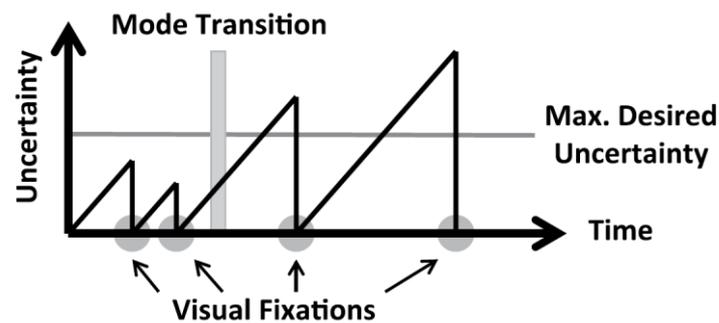


Abbildung 2-11 - Die wachsende Unsicherheit über einen Systemzustand zwischen Fixationen (Johnson et al., 2017)

## 2.2.2 MESSUNG

Wie bereits im Abschnitt 2.1.4 zur Messung von Aufmerksamkeit diskutiert wurde, gilt auch bei der Messung von Situationsbewusstsein eine Operationalisierung dieses Modells zu finden. Der fehlende wissenschaftliche Konsens über die Definition von Situationsbewusstsein bedingt, dass es eine Vielzahl unterschiedlicher Messmethoden gibt, deren Zuverlässigkeit und Validität nicht gesichert ist (Salmon et al., 2009). Dieses Kapitel gibt einen Überblick über Messmethoden und verwandte empirische Arbeiten. Basierend auf Salmon et al. (2006) und Salmon et al. (2009) werden die Methoden nach der folgenden Kategorisierung beschrieben: Bewertung mittels Abfrage von Informationen, Subjektive Bewertung und maschinelle Verhaltensbeobachtung.

### 2.2.2.1 Abfrage von Informationen

Im Zentrum der Situationsbewusstseinstheorie steht die Frage, ob sich ein Operateur den Zustand einer aufgabenrelevanten Information und deren Rolle für seine Aufgabe bewusst ist, weswegen eine direkte Abfrage dieser Information eine naheliegende Methode zur Messung ist. Voraussetzung für die Erfassung von Situationsbewusstsein mittels der Abfrage von Informationen ist es, die untersuchte Situation simulieren zu können. Dabei gibt es bei der Messung durch Abfrage von Informationen zwei Verfahren: Befragung bei gestoppter Simulation (engl. *freeze probe*) und laufender Simulation (engl. *real-time probe*). Die Antworten der Probanden werden mit dem aktuellen Systemzustand verglichen, um das vorhandene Situationsbewusstsein zu bewerten. Diese Art der Operationalisierung bedeutet auch, das Situationsbewusstsein gleichgesetzt wird mit den Informationen, die der Bediener in seinem Arbeitsgedächtnis hält und zurückmelden kann. Ferner setzt diese Methode eine normative Vorstellung von Situationsbewusstsein voraus, das heißt die Vorstellung, dass perfektes Situationsbewusstsein durch abfragbares Wissen darstellbar ist.

#### Befragung bei gestoppter Simulation

Bei dieser Methode wird die Simulation einer Aufgabensituation an zufällig gewählten Zeitpunkten „eingefroren“ und alle Anzeigeelemente ausgeblendet. Dann werden dem Bediener Fragen über den Zustand aufgabenrelevanter Informationen gestellt. Die *Situation Awareness Global Assessment Technique* (SAGAT) ist der bekannteste Vertreter dieser Art von Befragungen (Endsley, 1995a). Voraussetzung für den SAGAT ist eine ausreichende Analyse des Arbeitsprozesses und die Kenntnis über die situative Relevanz von Informationen (bspw. durch eine Aufgabenanalyse). Basierend auf dieser Analyse werden Abfragen vorbereitet, die auf relevante Informationen für die in der Aufgabenanalyse definierten Ziele und Entscheidungen abzielen. Dabei muss darauf geachtet werden, dass nicht mehrfach dieselben Informationen abgefragt werden, da die Probanden diesen sonst mehr

Aufmerksamkeit schenken als eventuell notwendig ist. Die Antworten der Versuchsteilnehmer werden mit dem tatsächlichen Zustand verglichen und damit ein Situationsbewusstseinswert festgelegt. Neben dem SAGAT, der in Untersuchungen von militärischen Luft-Luft-Bekämpfungen sowie der Luftverkehrskontrolle durchgeführt wurde, haben sich auch spezifische Methoden für die Domäne der Luftfahrt (Hogg et al., 1995) sowie der Kraftwerkssteuerung (Haus & Eyferth, 2003) etabliert, die relevante Informationen der jeweiligen Domäne integrieren.

Vorteil dieser Methodik ist die direkte und objektive Natur der Erfassung gegenüber subjektiven Fragebögen, die nach einem Versuch durchgeführt werden. Kritisiert wurde der SAGAT vor allem für seine störende Natur, da ein Einfrieren der Simulation und ein Ausblenden der Bedienoberfläche ein drastischer Eingriff in die Aufgabensituation ist, welche das Situationsbewusstsein negativ beeinflusst. Darüber hinaus ist auch offen, ob der SAGAT nicht eher das Erinnerungsvermögen der Probanden prüft. Ein weiterer Nachteil für die Vergleichbarkeit von Studien ist, dass die Abfragen im SAGAT immer domänenbezogen und meist auf die konkrete Aufgabensituation abgestimmt sind. Eine weiteren Kritikpunkt wird von Winter et al. (2019) diskutiert und ist in Abbildung 2-12 dargestellt.

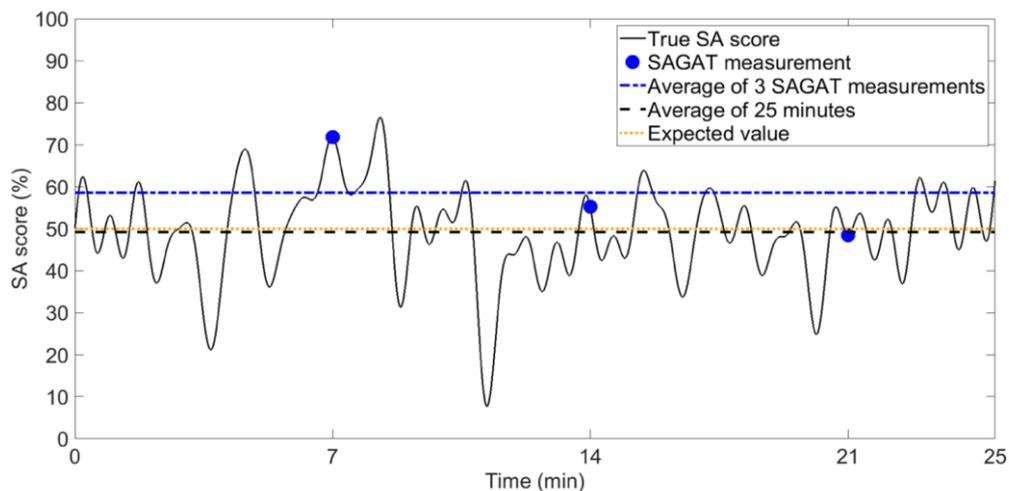


Abbildung 2-12 - Hypothetische Abbildung einer Situationsbewusstseinsmessung basierend auf der SAGAT (Winter et al., 2019)

Die Darstellung zeigt den hypothetischen Verlauf von Situationsbewusstsein in einem Versuch. Wenn der objektive SB-Wert so dynamisch ist und ein SAGAT nur zu wenigen diskreten Punkten prüft, kann man damit zu einem falschen Schluss über das Situationsbewusstsein des Bedieners gelangen.

### Befragung bei Durchführung

Bei dieser Methode werden Informationen während der Durchführung der untersuchten Aufgabe abgefragt, ohne sie zu unterbrechen. Als Maß für Situationsbewusstsein dient der Vergleich der Antworten mit dem Systemzustand. Eine zusätzliche Größe, die zur Bewertung beiträgt, ist die benötigte Zeit zur Beantwortung. Die Annahme ist hierbei, dass ein Bediener mit hohem Situationsbewusstsein die Fragen direkt beantworten kann, ohne sich die betreffenden Anzeigen noch einmal ansehen zu müssen. Diese Methode wurde als *Situation Present Assessment Method* (SPAM) bei Luftverkehrskontrolle eingesetzt. Dabei wurde ein Fluglotse über Telefon über den aktuellen Zustand des Luftraums befragt (bspw. „welches der Flugzeuge A und B fliegt höher?“). Die Dauer bis zur Annahme des Anrufes wurde als Arbeitsbeanspruchung interpretiert. Die Dauer bis zur Antwort der Frage sowie deren Korrektheit wurden als Indikator für das Situationsbewusstsein genommen (Durso et al., 1998). Vorteil dieser Methode ist, dass die Aufgabendurchführung nicht angehalten und

der Versuchsteilnehmer nicht aus seiner Situation gerissen wird. Dies ermöglicht den Einsatz in Anwendungen, in denen kein Pausieren der Aufgabe möglich ist. Nachteil ist, dass die Befragung zusätzliche Arbeitsbeanspruchung auslösen kann, welche den Bediener bei der Ausführung seiner Tätigkeit und damit auch der Bildung von SB stört. Außerdem kann das Prüfen bestimmter Informationen den Bediener auf Informationen hinweisen, denen er sonst keine Aufmerksamkeit geschenkt hätte. Dies verzerrt das Ergebnis der darauffolgenden Abfragen.

### 2.2.2.2 Subjektive Bewertung

Neben der Abfrage von Informationen, wurde Situationsbewusstsein in zahlreichen Studien mit Bewertung durch subjektive Einschätzung gemessen. Hierbei kann zwischen Bewertung auf Basis von Fragebögen oder durch Experten unterschieden werden.

### Selbstbewertung

Die Selbstbewertung von Situationsbewusstsein wird mittels Fragebögen erfasst, die nach Durchführung einer simulierten oder realen Arbeitssituation vom Bediener ausgefüllt werden. Einer der häufigsten eingesetzten Fragebogen dieser Art ist die *Situation Awareness Rating Technique* (engl. SART) (R.M. Taylor, 2017). Dabei werden zehn Dimensionen abgefragt, in denen ein Proband sich selbst auf einer Skala (1=niedrig, 7=hoch) bewerten soll. Der größte Vorteil dieser Methoden ist der geringe Aufwand der Durchführung. Gegenüber der Abfrage von Informationen greift die Selbstbewertung auch wesentlich weniger in die Durchführung einer Arbeitssituation ein und kann somit eine Wechselwirkung auf die situative Messgröße ausschließen. Ferner lassen sich die Ergebnisse von Studien, die SB mit dem standardisierten SART messen, besser vergleichen. Diesen Vorteilen gegenüber steht die grundlegende Problematik von Introspektion und Selbstbewertung des Bedieners und damit einhergehend einer geringeren Sensitivität des Messverfahrens.

### Bewertung durch Beobachter

Die Bewertung durch externe Beobachter basiert meist auf einer Bewertung des Situationsbewusstseins durch einen Experten. Dieser beobachtet das Verhalten eines Operateurs und schließt dabei auf sein Situationsbewusstsein. Vorteil ist hierbei, dass diese Methode nicht intrusiv ist, was sie gut für Feldstudien und Messungen in naturalistischer Arbeitsumgebung macht wie beispielsweise bei einer militärischen Geländeausbildung (Matthews & Beal, 2002). Außerdem wird die Problematik der fehlerbehafteten Introspektion vermieden. Nachteil der Methode ist, dass ein Experte meist nur beobachtbares Verhalten in aufgabenbezogenen Tätigkeiten bewerten kann, also nur in Situationen, in denen eine Handlung notwendig ist. Schwieriger ist die Abschätzung, ob sich der Proband bestimmter Systemzustände bewusst ist, wenn gerade keine Reaktion notwendig ist.

### 2.2.2.3 Maschinelle Bewertung

Eine Alternative zu den bisher beschriebenen Messverfahren ist die maschinelle Bewertung von Situationsbewusstsein durch ein technisches Messsystem. Bei diesem Ansatz wird angenommen, dass ein Zusammenhang zwischen einer beim Bediener messbaren Größe und dem Situationsbewusstsein existiert. Die Ansätze in der Forschungsliteratur unterscheiden sich bei der Auswahl der Messgröße und lassen sich in drei Kategorien einteilen: Kontextfrei, kontextbehaftet und situativ. Als Referenz für Situationsbewusstsein der im Folgenden diskutierten Studien werden meist Methoden wie Selbstbewertung mittels SART oder Befragung durch SAGAT oder SPAM verwendet.

## Kontextfreie Bewertung

Kontextfrei sind diejenigen Messgrößen, die sich unabhängig vom Mensch-Maschine-System messen lassen. Dabei ist man auf messbare Signale des menschlichen Körpers beschränkt. Typische Beispiele sind Herzrate oder EEG. T. Zhang et al. (2020) zeigen in einer Zusammenfassung der Forschungsliteratur mit Fokus auf den Zusammenhang von kontextfreien Messgrößen mit Situationsbewusstsein, dass Blickbewegungsmessung das verbreitetste Mittel ist gefolgt von kardiovaskularen Größen. Kontextfreie Metriken der Blickbewegungsmessung sind beispielsweise Blinzelraten, Pupillendilation oder Fixationsraten, wobei in keiner Studie ein signifikanter Zusammenhang zwischen diesen kontextfreien Größen und Situationsbewusstsein nachgewiesen werden konnte. Bei den kardiovaskularen Größen wie beispielsweise der mittleren Herzrate konnten in einzelnen Studien Korrelationen nachgewiesen werden (siehe beispielsweise Wei et al. (2013)). Die meisten Studien – wie beispielsweise die von Kunze et al. (2019) – fanden jedoch keinen Zusammenhang.

## Kontextbehaftete Bewertung

Kontextbehaftete Metriken können nur im Zusammenhang mit der Bedienung eines technischen Systems definiert werden. Sie beschreiben damit die Interaktion des Bedieners in seiner Arbeitssituation. So lässt sich beispielsweise die Fixationsrate auf ein Interessensgebiet nur messen, wenn diese Interessensgebiete sinnhaft im Sinne der Aufgabe definiert sind. Im Kontext der Luftfahrt wäre eine sinnvolle Aufteilung beispielsweise *Primary Flight Display* gegenüber *Navigationsdisplay*. Unter den kontextbehafteten Messgrößen dominiert die Blickbewegungsmessung. Zum Zeitpunkt dieser Arbeit konnten keine Studien gefunden werden, die alternative Ansätze wie Klickraten oder Verhaltensmuster im Zusammenhang mit Situationsbewusstsein untersucht haben.

Moore und Gugerty (2010) fanden in einer Studie mit Fluglotsen heraus, dass die relative Fixationszeit auf einem Interessensgebiet einen moderaten positiven Zusammenhang mit dem Ergebnis in einem dem SAGAT ähnlichen Bewertungsmaß für Situationsbewusstsein hat. Dabei waren die ausgewählten Interessensgebiete immer diejenigen, die in der Informationsabfrage geprüft worden sind. Darauf aufbauend untersuchten van de Merwe et al. (2012) den Zusammenhang zwischen Fixationsraten und Verweilzeiten auf Interessensgebieten mit dem subjektiven Bewertungsverfahren CARS. Die Ergebnisse einer Studie mit kleiner Probandenzahl deuteten auf einen positiven Zusammenhang dieser Größen hin. Im Kontext des Bauingenieurwesens wurde eine positive Korrelation zwischen selbstbewerteten Situationsbewusstsein und der Verweilzeit und Fixationsrate auf einem relevanten Interessensgebiet festgestellt, wobei kein Zusammenhang in anderen blickbewegungsbasierten Metriken nachgewiesen werden konnte (Hasanzadeh et al., 2016). In den T. Zhang et al. (2020) untersuchten 16 Studien über den Zusammenhang zwischen blickbewegungsbasierten Metriken und SB konnte der Zusammenhang zwischen Verweilzeit und SB nur für zwei von sechs Studien reproduziert werden. Die Fixationsrate konnte nur in einer von zwei Studien positiv mit SB korreliert werden. Diese Ergebnisse stellen die Validität der blickbewegungsbasierten Metriken als Messmethode für SB in Frage.

## Situative Bewertung

Situative Metriken bauen auf der kontextbehafteten Bewertung auf und ergänzen sie durch den Zustand der Aufgabensituation. Sie beinhalten damit einen Bezug zu den dynamischen Größen der Arbeitsumgebung.

Im Zusammenhang mit Situationsbewusstsein in der Führung unbemannter Systeme stellten Ratwani et al. (2010) einen Modell vor, dass sehr spezifisch für eine Aufgabenumgebung vorhersagen sollte,

ob der Operateur einen Fehler begehen wird. Dabei wurde basierend auf einem Experiment eine Regressionsmodell aufgestellt, das den Aufgabenzustand mit Fixationsraten kombiniert, um eine Fehlerwahrscheinlichkeit vorherzusagen. Dabei wurde kein konventionelles Maß zur Bewertung von Situationsbewusstsein herangezogen, sondern das Fehlermaß als Indikator für Situationsbewusstsein gewählt. Die Autoren begründen das damit, dass Situationsbewusstsein maßgeblich für die Leistung in der gestellten Aufgabe sei.

In einem ähnlichen Aufgabenkontext entwickelten Frische und Lüdtkte (2013) ein System zur Abschätzung von Situationsbewusstsein. Bemerkenswert an diesem Ansatz ist, dass es die einzige Arbeit ist, die die Semantik spezifischer Informationen in einer Aufgabenumgebung explizit modelliert. Damit wird Situationsbewusstsein nicht mehr rein als ein numerischer Wert angegeben, sondern auf bestimmte Zustände von Informationen im System bezogen. Abbildung 2-13 zeigt hier den konzeptionellen Aufbau mit Blickbewegungsmessung, Zustand der Bedienoberfläche und des Systems, welche mit einer Wissensbasis kombiniert werden. Ziel ist eine Abschätzung über fehlendes oder falsches Situationsbewusstsein.

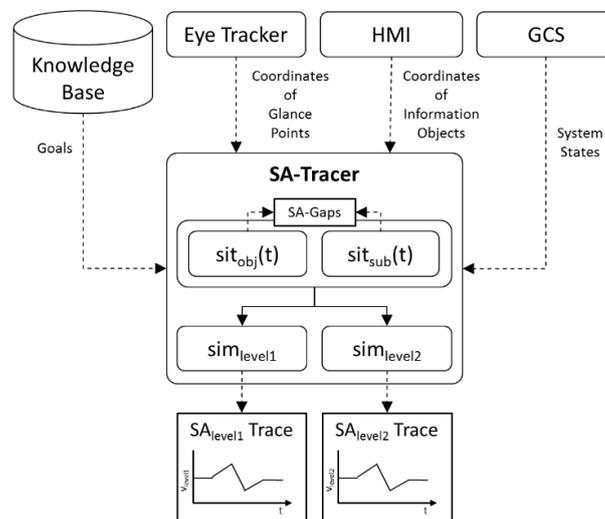


Abbildung 2-13 - Systemaufbau des technischen Messsystems aus Fortmann und Mengerlinghausen (2014)

Die Messkette funktionierte so, dass alle fixierten Elemente auf der Oberfläche nach ihrer Wahrnehmbarkeit modelliert wurden, welche durch die vier Bestandteile von Wickens SEEV-Modell definiert war. Wenn dann eine Fixation auf einem Element gemessen wurde und dieses Element über einem spezifizierten Wahrnehmungsgrenzwert lag, wurde diese Information als wahrgenommen gespeichert. Die Menge aller wahrgenommenen Informationen wurde dann kontinuierlich mit dem Systemzustand verglichen und damit eine Abschätzung für das Situationsbewusstsein des Bedieners generiert. Dabei beschränkten sich die Autoren auf skalare und boolesche Werte, die für Systemzustände in der Multi-UAV-Anwendungen standen (beispielsweise skalar = Treibstoff 0-100%, bool = Warnlampe an/aus). Das Modell beinhaltete eine Wahrnehmungs- und eine Aufgabenebene. Die Wahrnehmungsebene bezog sich auf die Informationen, die auf der Oberfläche dargestellt waren. Die Elemente auf dieser Ebene wurden mit Zielen der Aufgabensituation verknüpft, was die zweite Ebene darstellte. Eine dargestellte Information im System kann demnach für eine Aufgabe relevant sein, wenn sie mit dieser in der Wissensbasis verknüpft ist. Mit diesem Mechanismus wird eine Art des Wissens über das Arbeitsziel des Operateurs modelliert. In einer Aufgabenumgebung zur Führung von mehreren UAVs mit 10 Teilnehmern wurde die Vorhersageleistung dieser Messkette verglichen mit einer dem SAGAT ähnlichen Abfrage von Informationen nach einem kurzen Szenario von zwei

Minuten. Das System konnte - nach Optimierung des Wahrnehmbarkeitswerts - ein Matthews Korrelationskoeffizient von 0.43 erreichen mit 85.6% richtig positiver und 62.4% richtig negativer Klassifikation. Diese Arbeit zeigt somit den Vorteil eines semantischen Messsystems, das die Zustände der Aufgabe explizit modelliert.

Eine aktuelle Arbeit mit theoretischem Schwerpunkt beschäftigte sich mit der Validität konventioneller Messmethoden im Vergleich mit Blickbewegungsmessung (Winter et al., 2019). Aufbauend auf der bereits diskutierten Kritik am SAGAT, schlagen die Autoren vor, Blickbewegungen eines Bedieners mit Bezug zur Aufgabenumgebung auszuwerten und als Maß für Situationsbewusstsein zu nutzen. In einem Experiment mussten die Versuchsteilnehmer mehrere Messuhren überwachen und bei Anzeige innerhalb eines definierten Bereichs mit einem Knopfdruck reagieren. In dieser simplen Aufgabenumgebung wurde gezeigt, dass die Befragung basierend auf Einfrieren der Aufgabenumgebung viel schlechter mit der Leistung der Probanden korrelierte als einem einfachen Blickbewegungsmaß, das feststellte, ob die relevante Messuhr innerhalb eines kurzen Zeitraums ( $t < 0.5s$ ) vor notwendiger Reaktion fixiert wurde.

### 2.2.3 RESÜMEE

Der erste Teil dieses Abschnitts diskutiert verschiedene Modelle von Situationsbewusstsein, welche sich in ihrer zugrundeliegenden Perspektive des menschlichen Verhaltens im Umgang mit technischen Systemen unterscheiden. Der Überblick über die Rechnermodelle zeigt, dass sich hier die Modellvorstellung des Informationsverarbeitungsparadigmas durchgesetzt hat, da die Modelle des ökologischen Paradigmas keine greifbaren Größen im Sinne einer Operationalisierung bieten. Ein ähnliches Bild zeichnen die Messverfahren, die in Abfrage von Informationen, subjektive Bewertung sowie maschinelle Abschätzung unterteilt werden könnten.

Die Tabelle 2-1 stellt eine Bewertung der Methoden nach Genauigkeit, Aufwand, Transfer- und Onlinefähigkeit.

Kategorie	Methode	Genauigkeit	Aufwand	Transferfähig	Onlinefähig
<b>Abfrage</b>	Gestoppte Simulation	+	0	0	-
	Laufende Simulation	+	0	0	-
<b>Subjektive Bewertung</b>	Selbstbewertung	0	+	+	-
	Beobachter	0	+	0	-
<b>Maschinelle Bewertung</b>	Kontextfrei	-	0	+	+
	Kontextbehaftet	0	-	0	+
	Situativ	+	-	-	+

Tabelle 2-1 - Bewertung der Messmethoden von SB

Die Durchführung etablierter Abfragemethoden haben einen moderaten Aufwand durch Simulation der Aufgabenumgebung, können aber eine hohe Genauigkeit erzielen, welche durch eine Vielzahl von Studien bestätigt wurde. Aufgabenspezifische Ergebnisse lassen sich nur eingeschränkt übertragen, da die Abfragen meist speziell auf die Aufgabensituation zugeschnitten sind. Subjektive Bewertungen sind besonders einfach durchzuführen und durch eine mögliche Standardisierung (bspw. SART) auch vergleichbar und transferfähig. Sie leiden jedoch unter den Problemen subjektiver Bewertung. Zur onlinefähigen Verwendung in einem adaptiven System eignen sich von den gesichteten Methoden besonders die maschinellen Messverfahren. Dabei sind die Arbeiten über kontextfreie Messgrößen nicht vielversprechend, da nur vereinzelt Zusammenhänge zwischen Messgröße und SB oder Leistung nachgewiesen werden konnten. Darüber hinaus ist es auch grundsätzlich fraglich, ob kontextfreie

Größen ein domänenbehaftetes Konstrukt wie Situationsbewusstsein ausreichend beschreiben können. Die kontextbehafteten Messgrößen sind hier vielversprechender und können, kombiniert mit dem Systemzustand in situativen Messgrößen, konkrete Aussagen darüber treffen, in welchen Systemzuständen das Bewusstsein falsch oder nicht vorhanden ist. Zuletzt zeigt auch die Studie von Winter et al. (2019) das diese Art der Messung Vorteile gegenüber den konventionellen Verfahren in der Vorhersage der menschlichen Leistung bietet. Ein Problem der situativen Verfahren liegt im Umsetzungsaufwand aber vor allem im Bereich des Transfer, da diese Ansätze meist sehr spezifisch auf die gewählte Aufgabenumgebung zugeschnitten sind (beispielsweise bei Fortmann und Lüdtker (2013)). Diese Problematik wurde bei der Methodenentwicklung dieser Arbeit aufgegriffen, um dort eine Verbesserung zu erzielen.

## 2.3 SITUATIONSBEWUSSTSEIN IM KONTEXT MENSCHLICHER FEHLEISTUNG

Bei der Analyse von 143 Flugunfällen kommen D. G. Jones und Endsley (1996) zu dem Schluss, dass bei 111 der Unfälle Fehler auf einer der drei Situationsbewusstseinsebenen passierten. Ein ähnliches Bild zeichnen Kharoufah et al. (2018), die bei der Analyse von 200 Flugunfällen fehlendes Situationsbewusstsein als häufigsten menschlichen Fehler anführen. Das vorliegende Kapitel soll klären, welche Art von Fehlern durch fehlendes Situationsbewusstsein begünstigt werden. Hierfür wird zunächst mögliche Klassifikationen von menschlichen Fehlern vorgestellt. Darauf aufbauend werden fehlerbegünstigende Phänomene im Umgang mit technischen Systemen diskutiert.

### 2.3.1 TAXONOMIE DER MENSCHLICHEN FEHLER

Um zu ergründen, welche Arten von Fehlern im Zusammenhang mit SB stehen, ist es nützlich das Fehlermodell von Reason (1990b) heranzuziehen, in dem Fehler grundsätzlich definiert wird als ein

[...] GENERIC TERM TO ENCOMPASS ALL THOSE OCCASSIONS IN WHICH A PLANNED SEQUENCE OF MENTAL OR PHYSICAL ACTIVITIES FAILS TO ACHIEVE ITS INTENDED OUTCOME [...].

Fehler können in unterschiedliche Kategorien geordnet werden, welche sich auf Rasmussen's Modell der menschlichen Leistung stützt (Rasmussen, 1983). Zunächst wird unterschieden, ob ein Fehler bei der Ausführung der richtigen Aktion entstand oder ob die geplante Absicht des Benutzers bereits fehlerhaft war. Ersteres wird als Ausrutscher und Versehen (engl. *Slips and Lapses*) bezeichnet. Ein Ausrutscher (engl. *Slips*) ist ein Fehler, der bei der Durchführung einer planmäßig richtigen Aktion passiert. Ausrutscher werden beispielsweise begünstigt durch Überlastung oder zeitlich falsch platzierte Aufmerksamkeit auf die ausgeführte Tätigkeit. Besonders bei routinierten Aufgaben, deren Ausführung eher einem gesteuerten als einem geregelten Vorgehen entspricht, besteht das Risiko, dass das Ergebnis der Aktion nicht adäquat überprüft wird. Ein Versehen (engl. *Lapse*) ist dagegen weniger offensichtlich und bezieht sich auf Fehler des Gedächtnisses, bei dem beispielsweise eine Aufgabe nicht mehr an der richtigen Stelle wiederaufgenommen wird, nachdem sie unterbrochen wurde.

Den Ausrutschern und Versehen gegenüber stehen Fehler, die auf der Ebene der Situationsinterpretation und Handlungsplanung liegen. Diese werden als Irrtümer (engl. *Mistakes*) bezeichnet und unterschieden in regel- und wissensbasierte Irrtümer. Erstere passieren in Situationen, in denen ein Mensch vertraute Probleme mithilfe von gelernten Regeln lösen will. In diesem Fall resultiert ein Fehler aus einer falschen Klassifikation der Situation und führt entweder zur Anwendung einer falschen Regel oder zur falschen Anwendung einer richtigen Regel. Während falsche Regeln durch fehlerhaftes oder fehlendes Training und Erfahrung vorliegen können, kann die falsche

Anwendung einer richtigen Regel das Resultat eines falschen Situationsbilds sein, welches durch eine ungünstige Verteilung von Aufmerksamkeit auf Systeminformationen verursacht ist.

Die wissensbasierte Ebene spielt dann eine Rolle, wenn ein Mensch unbekannten Situationen begegnet und er mittels eines bewussten, kognitiven Planungsprozesses eine Entscheidung treffen muss. Fehler in dieser Ebene treten durch die Limitierung der menschlichen Informationsverarbeitungsressourcen auf oder durch unvollständiges oder fehlerhaftes Wissen. Ähnlich wie bei der fehlerhaften Anwendung von Regeln, könnte auch hier eine Selektion der falschen Systeminformationen wissensbasierte Irrtümer begünstigen.

Die Fehleranalyse in der Luftfahrt integriert neben dem menschlichen Fehler, meist noch institutionelle Fehler sowie andere Aspekte, die zu Luftfahrnfällen beitragen. So hat Reason (1990a) seine Taxonomie der menschlichen Fehler in das *Schweizer-Käse Modell* integriert. Fehlerhafte Verteilung von Aufmerksamkeit taucht hier nur als einer von vielen Gründen für menschliche Fehler auf. Sie spielt aber eine größere Rolle in der verbreiteten Unfallanalysemethode *Human Factors Analysis and Classification System* (Wiegmann & Shappell, 2003). Die Autoren schreiben fehlerhafte Verteilung von Aufmerksamkeit eine zentrale Rolle bei menschlichen Fehlern im Flugzeugcockpit zu:

“IN FACT, ATTENTION FAILURES HAVE BEEN LINKED TO MANY SKILL-BASED ERRORS SUCH AS THE BREAKDOWN IN VISUAL SCAN PATTERNS, TASK FIXATION, THE INADVERTENT ACTIVATION OF CONTROLS, AND THE MISORDERING OF STEPS IN A PROCEDURE, AMONG OTHERS.” (WIEGMANN & SHAPPELL, 2003, S. 51)

Diese Einschätzung ignoriert aber, dass eine falsche Verteilung von Aufmerksamkeit nicht nur für fertigkeitbasierte Fehler verantwortlich ist, sondern auch der Grund für Irrtümer und falsche Entscheidungen sein kann. D. G. Jones und Endsley (1996) stellen ein alternatives Fehlermodell vor, das auf ihrem Ebenen-Modell von SB basiert. Sie argumentieren, dass die meisten Fehler erfahrener Bediener aus der Kategorie der Irrtümer durch niedriges SB begründet sind. Sie unterscheiden hier in Fehler auf den drei Ebenen des SBs:

- Level 1: Information nicht wahrgenommen oder falsch wahrgenommen:
  - Information nicht verfügbar
  - Information nicht bemerkt
  - Information nicht überwacht
  - Information falsch wahrgenommen
  - Information vergessen
- Level 2: Fehlerhafte Integration und Verständnis der Information:
  - Fehlerhaftes mentales Modell
  - Zu hohes Vertrauen
- Level 3: Falsche Projektion des zukünftigen Systemzustands
  - Fehlerhaftes mentales Modell

Aber wie genau kommt es zu diesen Fehlern? Einige Forschungsarbeiten haben in der Interaktion von Menschen und Automation fehlerbegünstigende Phänomene identifiziert, die auf das SB des Bedieners negativen Einfluss haben.

### 2.3.2 FEHLERBEGÜNSTIGENDE PHÄNOMENE

Bei der Bedienung von Maschinen in Arbeitssituationen sind verschiedene Phänomene bekannt, die menschliche Fehler begünstigen. Einige sind durch die Aufgabensituation selbst verursacht, wie das

Wachsamkeitsdekrement, andere durch die Gestaltung des technischen Systems und der verwendeten Automationsgrade, wie beispielsweise die Gleichgültigkeit bei der überwachenden Kontrolle. Im Folgenden werden vier dieser Phänomene diskutiert.

### Wachsamkeitsdekrement

Viele Arbeitssituationen verlangen eine hohe Wachsamkeit (engl. *Vigilance*) und die Fähigkeit, Aufmerksamkeit über einen längeren Zeitraum so zu fokussieren, dass man relevante Stimuli zeitnah erkennt. Mit steigendem Einsatz von Automation in operationellen Arbeitsplätzen ist Wachsamkeit und die Tätigkeit der Überwachung eine entscheidende Komponente der menschlichen Leistung (Sheridan, 1992). Das Wachsamkeitsdekrement beschreibt die sinkende menschliche Leistung bei einer monotonen Überwachungsaufgabe, wobei Leistung stets auf die Detektion von aufgabenrelevanten Stimuli bezogen wird (Warm et al., 2008). Das Leistungsdekrement kann bei hoher Aufgabenbelastung bereits ab fünf Minuten beginnen (Helton et al., 2007) und konnte sowohl im Labor als auch in einer realistischen Arbeitsumgebung nachgewiesen werden. In der wissenschaftlichen Literatur gibt es zwei konträre Erklärungen für das Wachsamkeitsdekrement: Eine Theorie besagt, dass das Dekrement durch die hohe Beanspruchung der Wachsamkeitsaufgabe ausgelöst wird und der Bediener mit der Zeit kognitive Ressourcen verliert (Warm et al., 2008). Die konkurrierende Theorie argumentiert gegensätzlich und besagt, dass die zu niedrige Beanspruchung für ein Wandern der Gedanken verantwortlich ist, welche die Aufmerksamkeit des Bedieners beanspruchen und von der primären Aufgabe ablenken (McVay & Kane, 2012). Thomson et al. (2015) vereinen beide Theorien derart, dass das Wachsamkeitsdekrement durch eine abnehmende *exekutive Aufmerksamkeitskontrolle* entsteht, welche das Wandern der Gedanken begünstigt. Diese Kontrolle nimmt ab, weil die geringe Frequenz der Stimuli vom Menschen gelernt wird und es immer mehr Willenskraft kostet, bewusst gegen das gelernte Überwachungsverhalten zu arbeiten.

### Aufmerksamkeitstunnel

Ein weiteres Phänomen ist der *Aufmerksamkeitstunnel* (engl. *attentional tunneling*). Hierbei handelt es sich um die unverhältnismäßig lange und fokussierte Allokation von Aufmerksamkeit auf einen Informationskanal, eine Diagnose oder ein Ziel. Diese überproportionale Allokation führt zur Vernachlässigung anderer Informationen, Diagnosen oder Ziele (Wickens, 2005). Die übermäßige Fixierung führt meist zu einer schlechteren Bearbeitung anderer Aufgaben des Operateurs. Dieses Phänomen wurde unter anderem als Ursache für *Controlled-Flight-Into-Terrain-Unfälle* identifiziert (Shappell & Wiegmann, 1998).

Im Kontext von Luftfahrtanwendungen, wurde der Aufmerksamkeitstunnel experimentell zunächst im Zusammenhang mit synthetischen Sichtsystemen (*Synthetic Vision System, SVS*) nachgewiesen (Wickens et al., 2004). Hierbei flogen 14 Piloten in einem Simulator Flugpfade entlang in unterschiedlichen Konfigurationen des SVS. Gegenüber der Kontrollbedingung, erhöhte das SVS eindeutig die Leistung in der Flugpfadverfolgung. Es konnte aber gezeigt werden, dass unerwartete Situationen oder Fehler, die nur in der Außensicht sichtbar waren und nicht im SVS nicht abgebildet sind, vom Piloten mit höherer Wahrscheinlichkeit nicht erkannt werden. Diese Ergebnisse wurden von Alexander et al. (2005) reproduziert, wobei keine Kontrollbedingung untersucht wurde. Laut der Autoren ist der Grund hierfür ein zu hoher Aufmerksamkeitsanteil auf Elemente der künstlichen Sicht, während die reale Außenwelt vernachlässigt wird. Dieser Effekt verstärkt sich sogar, wenn zur unerwarteten Situation noch ein Systemfehler hinzukommt und der Pilot sich in der Bearbeitung dieses Fehlers befindet. Jedoch ist es nicht klar, ob man diesen Bestand wirklich einem Aufmerksamkeitstunnel zusprechen sollte und nicht einer erhöhten Arbeitsbelastung oder einer höheren Priorität von Fehlermanagement.

Eine Studie beschäftigte sich in diesem Kontext mit der Erkennung des Aufmerksamkeitstunnels mittels maschinellen Lernverfahren, die laut der Autoren nützlich wäre, um die Schnittstelle des Benutzers anzupassen und damit den Benutzer aus dem Tunnel zu führen (Regis et al., 2014). Hierfür machten sie sich in einem Experiment zunutze, dass ein Nutzer besonders oft in einen Aufmerksamkeitstunnel gerät, wenn er unter hoher Arbeitsbelastung steht. In einem Versuch musste ein Nutzer den Sensor eines Roboters mit einem Joystick steuern. Während einer besonders kritischen Phase, wurde ein Fehler ausgelöst, der den Roboter zu seiner Ladestation fahren ließ, sofern der Benutzer dies nicht durch seinen Joystick verhinderte. Die Autoren evaluierten mit diesem Experiment eine Gegenmaßnahme, die darin bestand, das vom Benutzer fokussierte Display kurzzeitig auszublenden um ihn auf den Fehlerzustand aufmerksam zu machen. Verglichen mit einer Kontrollgruppe ohne Gegenmaßnahme, verringerte diese Anpassung des Displays den Tunnel um rund 67%.

Obwohl sich in dieser Studie dem Aufmerksamkeitstunnel als kognitives Phänomen gewidmet und dieser Zustand sogar maschinell klassifiziert wird, wird über die Existenz des Aufmerksamkeitstunnels noch gestritten, da die Fehler auch mit anderen Effekten erklärt werden können (Wickens, 2005). Dass die Piloten mit SVS nicht die unerwarteten Situationen erkannt haben, weil diese nicht im SVS erkennbar waren, könnte auch auf ein zu hohes Vertrauen in das Sichtsystem zurückzuführen und damit mit einer Verteilungsverzerrung der Aufmerksamkeit (engl. *Attentional Bias*) zu erklären sein. Tatsächlich wurde eine zweite unerwartete Situation - ein zu umfliegendes Hindernis - in der Studie von Wickens et al. (2004) von allen Piloten bemerkt, da das Hindernis im Display des SVS zu sehen war. Die Frage ist also, ob man nicht bemerkte Informationen einem „Aufmerksamkeitstunnel“ zuordnet nur, weil die Piloten davon ausgingen alle relevanten Informationen in der künstlichen Sicht dargestellt zu bekommen. Die geringe Erfahrung in der Experimentalumgebung könnte ein weiterer Faktor sein. Es überrascht zudem auch nicht, dass diese Fehler öfter auftreten, wenn der Pilot mit einem Fehlerzustand im Cockpit beschäftigt ist. Fehlerzustände verursachen Stress und damit auch einen Aufmerksamkeitstunnel, in dem sich der Pilot nur noch mit der Beseitigung des Fehlers beschäftigt und andere Aufgaben vernachlässigt.

### Gleichgültigkeit bei der überwachenden Kontrolle

Das Phänomen der Gleichgültigkeit (engl. *Complacency*) wurde zunächst durch Wiener (1981) als psychologischer Zustand beschrieben, der durch ein geringes Maß an Misstrauen gegenüber dem Zustand eines technischen Systems charakterisiert ist. Weitere Studien ergaben zwar keinen Konsens über die genaue Definition von Gleichgültigkeit, jedoch beinhalten die meisten Definition die folgenden drei Aspekte: Erstens, das Phänomen der Überwachung tritt bei der Kontrolle eines technischen Systems durch einen Menschen auf (engl. *Monitoring*). Zweitens, die Frequenz und Präzision dieser Überwachung ist nicht optimal, was, drittens, negative Auswirkungen auf die Leistung des Mensch-Maschine-Gesamtsystems hat (Parasuraman & Manzey, 2010). In der Regel ist die negative Auswirkung hier das Übersehen einer Systemstörung oder -warnung und damit das Unterlassen einer Reaktion.

Parasuraman et al. (1993) führte eine Studie mit der *Multi-Attribute Task Battery* (MATB) durch, bei der eine Überwachungsaufgabe unterstützt wurde durch ein nicht vollständig zuverlässiges System. Dabei konnte nachgewiesen werden, dass sowohl die Aufgabenbelastung als auch die Zuverlässigkeit eines Unterstützungssystems Auswirkungen auf die Erkennungsrate von Fehlern hatte. Während die Erkennung von Fehlern in der Single-Task Kontrollbedingung fast perfekt war, führte eine hohe Mehrfachaufgabenlast zu einer niedrigeren Erkennungsrate, was darauf hindeutet, dass die Versuchsteilnehmer sich unter hoher Belastung mehr auf das automatisierte System verließen.

Interessanterweise hatte der Betrag einer konstanten Zuverlässigkeitsrate selbst nicht so große Auswirkungen auf die Erkennungsrate, wenn man sie mit einer variablen Zuverlässigkeitsrate verglich. In einer nachfolgenden Studie ließen Molloy und Parasuraman (1996) das Unterstützungssystem nur noch einmal fehlschlagen, was dazu führte, dass der Gleichgültigkeitseffekt stärker zutage trat. Auch in dieser Studie konnte gezeigt werden, dass sich die Aufgabenlast negativ auf die Erkennungsrate auswirkt. Darüber hinaus wurde gezeigt, dass es einen Unterschied macht, ob das System zu Beginn des Versuchs fehlschlägt oder erst am Ende. Dies ist einleuchtend und könnte darauf hindeuten, dass die Bediener Vertrauen in das automatisierte System gewannen und damit ihre Aufmerksamkeit auf eine Nebenaufgabe richteten, was als rationales Verhalten bezeichnet werden kann.

### Voreingenommenheit bei Entscheidungsunterstützungssystemen

Eines mit der Gleichgültigkeit verwandtes Phänomen ist die Voreingenommenheit im Zusammenhang mit automatisierten Entscheidungsunterstützungssystemen (engl. *automation bias*). Diese sind dafür da, die Entscheidungsfindung eines Operators zu unterstützen, sei es durch Benachrichtigungen über kritische Zustandsänderungen des Systems oder bereits einer Handlungsauswahl oder -empfehlung. Letzteres stellt den typischen Fall der automations-induzierten Voreingenommenheit dar, bei dem Vorschläge des Systems ohne kritisches Hinterfragen durchgeführt werden. Wickens et al. (2015) schlägt vor, die Gleichgültigkeit und Voreingenommenheit insofern abzugrenzen, dass sich Gleichgültigkeit in der Vernachlässigung der Überwachungsaufgabe niederschlägt und damit eher zu Unterlassungen einer Handlung führen (engl. *error of omission*), während die Voreingenommenheit zur Durchführung falscher Handlungen führt (engl. *error of commission*).

Empirisch konnte die durch Unterstützungssysteme induzierte Voreingenommenheit beispielweise von Layton et al. (1994) nachgewiesen werden. In einem Experiment mit Piloten wurde die Qualität einer Flugplanung mit verschiedenen Unterstützungsgraden evaluiert. Bei einem niedrigen Grad der Unterstützung waren die meisten Entscheidungen dem Piloten überlassen wobei spezifizierte Pläne automatisiert evaluiert wurden. Dem gegenüber stand der hohe Unterstützungsgrad, bei dem ein System einen Plan spezifiziert und empfahl. Piloten, die viel Unterstützung bekamen, betrieben wesentlich weniger Aufwand alternative Pläne zu evaluieren und waren auch anfällig dafür, fehlerhafte Pläne zu akzeptieren. Ähnliche Ergebnisse wurden in einer anderen Studie erzielt, die unterschiedliche Grade der Unterstützung zur Reaktion auf Tragflächenvereisung im Flug untersuchte (Sarter et al., 2001). Dabei konnte festgestellt werden, dass die Entscheidungsunterstützung zwar Leistung steigerte und Fehler reduzierte, aber im Falle einer falschen Empfehlung des Systems sogar widersprüchliche Informationen ignoriert wurden und der falschen Empfehlung folgegeleistet wurde. In einer etwas aktuelleren Studie von Manzey et al. (2012) wurde genauer beleuchtet, inwiefern unterschiedliche Grade der Unterstützung das Auftreten von Voreingenommenheit begünstigen. Dabei wurde festgestellt, dass der Unterstützungsgrad keinen Einfluss auf das Verfolgen falscher Vorschläge hatte. In diesem Zusammenhang konnte aber beobachtet werden, dass die Teilnehmer nach der Entscheidungsempfehlung zwar relevante Systemparameter überprüften aber wohl nicht kognitiv verarbeiteten und trotz widersprüchlicher Systemzustände der Empfehlung folgten. Dies könnte dem „*looked-but-failed-to-see*“-Phänomen zugeschrieben werden, bei dem ein Operateur relevante Informationen zwar fixiert aber nicht verarbeitet beziehungsweise richtig interpretiert. Dies passt auch zu den Ergebnissen von Sarter et al. (2007), bei dem Piloten trotz Fixationen auf die Anzeige des FMS-Modus einen Fehler bei der Betriebsartwahl begangen haben.

### 2.3.3 RESÜMEE

Zu Beginn dieses Abschnitts wird eine Taxonomie von menschlichen Fehlern eingeführt und im Zusammenhang mit dem Ebenen-Modell von SB besprochen. Im Anschluss daran werden vier

Phänomene der Mensch-Automations-Interaktion beschrieben, die einen negativen Einfluss auf die Aufmerksamkeitsverteilung des Menschen haben.

Die Konsequenz der vier diskutierten negativen Phänomene im menschlichen Umgang mit technischen Systemen ist eine falsche Verteilung von Aufmerksamkeit bei der Bedienung. Beim Wachsamkeitsdekrement und beim Aufmerksamkeitstunnel werden Informationen zugunsten anderer Tätigkeiten vernachlässigt. Bei der Gleichgültigkeit und Voreingenommenheit werden Systeminformationen nicht überwacht beziehungsweise überprüft. Dies führt in dynamischen Systemen unvermeidlich dazu, dass der Bediener kein richtiges Situationsbild mehr hat und damit Fehler aus der Kategorie Irrtümer wahrscheinlicher werden.

## 2.4 KOOPERATIVE AUTOMATION

Nachdem das vorhergehende Kapitel darlegt, wie Fehler im Umgang mit komplexer Automation entstehen können, soll sich der vorliegende Abschnitt mit der Frage beschäftigen, wie diese mit technischen Mitteln verhindert werden können. Eine (kognitions-)ergonomische Gestaltung des Arbeitsplatzes und die adäquate Ausbildung von Bedienern können das Risiko von Bedienfehlern reduzieren. Diese beiden Ansätze zur Fehlerreduktion - Systemgestaltung und Ausbildung - sind aus der Perspektive der Arbeitssituation jedoch statisch. Statisch, weil sowohl die Arbeitsplatzgestaltung als auch die Ausbildung des Bedieners bereits durchgeführt wurden und in diesem Moment nicht mehr geändert werden. Ein technischer Ansatz, der diese statische Natur aufbricht, ist die der kooperativen Automation, die ähnlich wie ein menschlicher Kooperationspartner in der Lage sein soll, menschliche Fehler zu entdecken, zu antizipieren und zu verhindern.

Diese Idee stützt sich auf der Theorie von Onken und Schulte (2010), wonach ein Bediener in zwei unterschiedlichen Verhältnissen mit intelligenter Automation interagieren kann. Der erste Modus ist hierarchisch, bei dem ein Bediener der Automation Aufgaben delegiert, welche von dieser ausgeführt werden. Im Umgang mit Automation im Flugzeugcockpit wäre das beispielsweise die Programmierung eines Autopiloten zum Halten einer Flughöhe. Der zweite Modus ist kooperativ, bei dem ein Bediener gemeinsam mit der Automation am Erreichen des Arbeitsziels arbeitet. Als Voraussetzung hierfür postulieren die Autoren, dass das technische System über kognitive Fähigkeiten verfügt, die es in die Lage versetzen, das Arbeitsziel des Mensch-Maschine-Systems zu interpretieren. Sobald die Automation das zu einem gewissen Grad kann, wird sie als kognitive Automation bezeichnet. Mit der Fähigkeit, das Arbeitsziel eigenständig zu verfolgen, kann sie den Piloten damit auch bei dieser Tätigkeit unterstützen.

Um zu verhindern, dass sich das Assistenzsystem und der Bediener im kooperativen Verhältnis stören und die Leistung des Gesamtsystems sinkt, ist die Frage zentral, wie eine gute Zusammenarbeit zwischen Menschen und kooperativer Automation erreicht werden kann. Dieser Modus der Zusammenarbeit wird in der Forschungsliteratur gemeinhin als *Human-Autonomy-Teaming (HAT)* bezeichnet (O'Neill et al., 2022). Inspiration in dieser Frage bietet der Blick auf menschliche Kooperation in Arbeitssituationen. Beispielsweise argumentieren Johnson et al. (2014), dass eine optimale Strategie zur dynamischen Aufgabenverteilung zwischen Bediener und Automation die der Dynamik zwischen Pilot-Flying und Pilot-Monitoring entspricht.

“BOTH THE PF [(PILOT FLYING)] AND PM [(PILOT MONITORING)] KNOW WHAT TASKS THEY ARE RESPONSIBLE FOR AT ANY POINT IN THE FLIGHT. WHEN THE PHASE OF FLIGHT CHANGES, THE TASKS PERFORMED BY EACH PILOT ALSO CHANGE. IN THIS SENSE, A CHANGE IN THE PHASE OF FLIGHT IS A CRITICAL EVENT THAT INHERENTLY TRIGGERS DYNAMIC TASK ALLOCATION. AND

WHEN THIS DYNAMIC TASK ALLOCATION OCCURS, THE PF DOES NOT HAVE TO TELL THE PM WHAT TO DO.” (JOHNSON ET AL., 2014)

Die Aufgabenteilung zwischen Assistenzsystem und Bediener ist nur eine der vielzähligen Gestaltungsmöglichkeiten von Zusammenarbeit. Es zeigt sich aber an diesem Beispiel, dass die Vision der kooperativen Automation durch eine Adaption des technischen Systems an den Kontext der Arbeitssituation umgesetzt werden könnte. Solche adaptiven System werden von Feigh et al. (2012) als technische Komponenten in Mensch-Maschine-Systems definiert, die ihr Verhalten an die sich ändernden Bedürfnisse des Bedieners anpassen. Zusammengefasst haben adaptive Systeme zwei zentrale Eigenschaften:

- Adaptive Systeme verändern ihr Verhalten basierend auf einer a-priori definierten Strategie, **der Adaptionstrategie**, welche die Arbeitssituation des Bedieners mit einbezieht.
- Zur Umsetzung der Adaptionstrategien werden **Kontextinformationen** genutzt, also Informationen, die die Arbeitssituation beschreiben und über das hinausgehen, was für die unterstützte Aufgabe notwendig wäre.

Diese beiden Eigenschaften dienen der Abgrenzung von adaptiven Systemen zur konventionellen Automation. Als Illustration hierfür soll ein Abstandregeltempomat (engl. *Adaptive Cruise Control*) dienen. Diese Funktion regelt die Geschwindigkeit des Autos so, dass der Abstand des vorausfahrenden Fahrzeugs miteinbezogen und ein Mindestabstand basierend auf der aktuellen Geschwindigkeit eingehalten wird. Obwohl *adaptiv* sogar in der englischen Bezeichnung des Systems steckt, ist die Grundfunktionalität hinsichtlich der hier verwendeten Definition nicht adaptiv, da sie keinerlei Adaptivität bezüglich der Situation des Bedieners aufweist. Wie könnte nach dieser Definition ein adaptiver Abstandregeltempomat aussehen? Dieser könnte beispielsweise den Mindestabstand nicht nur mithilfe der aktuellen Geschwindigkeit festlegen, sondern auch von der Vertrautheit des Bedieners mit dem System abhängig machen. Vertrautheit könnte beispielsweise durch Bedienzeit quantifiziert werden.<sup>3</sup>

#### 2.4.1 AUFBAU ADAPTIVER SYSTEME

Wie die Definition adaptiver Systeme bereits erkennen lässt, spielt die Analyse des Kontexts eine zentrale Rolle für die Umsetzung von adaptiven Strategien. Die generelle Struktur eines solchen adaptiven Systems ist in Abbildung 2-14 dargestellt. Dabei fließen vier verschiedene Klassen von Informationen in die Analyse des Bedienerkontexts: der Zustand des Bedieners, der Mission, des technischen Systems sowie der Welt. Basierend auf der Analyse dieses Kontexts wird dann gegebenenfalls eine Adaption generiert, welche sich auf die Schnittstelle zum Bediener oder auf die Maschine selbst auswirken kann. Die Gestaltung der Kontextanalyse und der Adaptionstrategie hat viele Freiheitsgrade, weshalb die Evaluierung des gesamten Mensch-Maschine-Systems unabdingbar ist. Da die Adaption basierend auf Bedienerbeobachtung im Zentrum der vorliegenden Arbeit liegt, werden im Folgenden Forschungsarbeiten aus diesem Kontext vorgestellt.

---

<sup>3</sup> Ob dies eine sinnvolle Adaption ist, steht auf einem anderen Blatt.

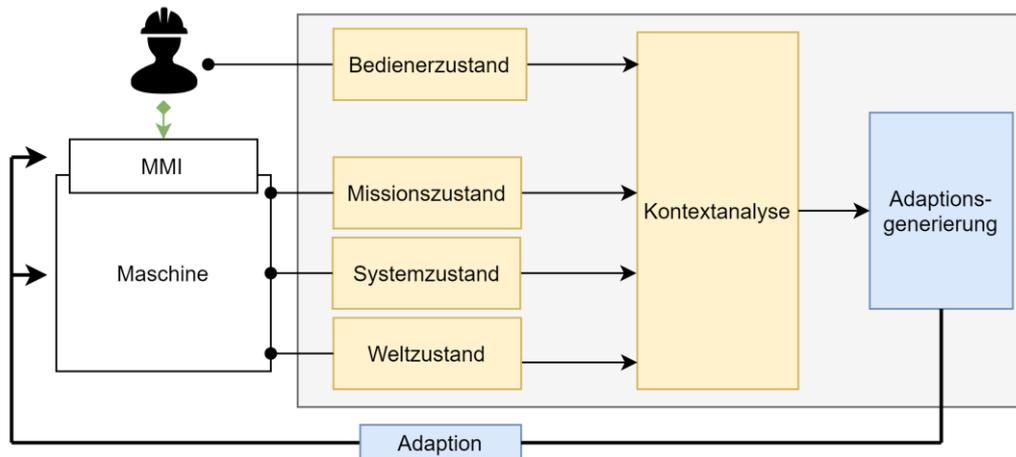


Abbildung 2-14 - Generischer Aufbau eines adaptiven Systems adaptiert von Feigh et al. (2012)

## 2.4.2 ADAPTION AUF BASIS VON BEDIENERBEOBACHTUNG

In diesem Abschnitt werden Studien mit unterschiedlichen Adaptionstrategien vorgestellt. Dabei wird ein besonderes Augenmerk auf diejenigen gelegt, die sich an den kognitiven Zustand des Bedieners anpassen. Hierbei wurden die Studien danach sortiert, welches kognitive Konstrukt als Grundlage für die Umsetzung dieser Strategie herangezogen wurde.

### Beanspruchungsadaptive Systeme

Frühe Studien mit adaptiven Ansätzen entstanden aufgrund von Arbeitsbelastungsproblemen bei der Bedienung von Automatisierung. Es wurde deshalb als sinnvoll erachtet, die Arbeitsbelastung durch Anpassung an die Automatisierung zu verringern (Kaber & Endsley, 2004). Zur Messung der Arbeitsbeanspruchung konzentrieren sich die meisten neueren Studien auf die Verwendung physiologischer Sensoren wie EEG. So verwendeten Wilson und Russell (2007) eine Kombination aus der Messung der EEG und des Elektrokardiogramms (EKG) zur Klassifizierung von hoher und niedriger Arbeitsbeanspruchung. Sie führten ein Experiment durch, bei dem die Teilnehmer vier unbemannte Fahrzeuge unter den folgenden Bedingungen bedienten: ohne Hilfe, mit adaptiver Hilfe und mit zufälliger Hilfe. Der Hilfsmechanismus hatte zwei Ziele: Erstens verringerte er den Zeitdruck und zweitens verbesserte er die Integration von Informationen auf dem Display. Die Ergebnisse zeigten, dass der Assistenz-Mechanismus die Leistungsmessungen um 50 % im Vergleich zu keiner Assistenz verbesserte. Außerdem verbesserte sich die Leistung im Vergleich zu einer zufälligen Hilfestellung. In einer neueren Studie wurde ein adaptives Hilfesystem für eine ATC-Aufgabe implementiert (Aricò et al., 2016). Die Autoren klassifizierten die Arbeitsbeanspruchung der Teilnehmer anhand von EEG-Messungen. Wenn die Arbeitsbelastung hoch war, wurden vier verschiedene Anpassungen vorgenommen. Ihre Ergebnisse zeigten, dass die Aufgabenleistung verbessert wurde, wenn die adaptive Unterstützung aktiv war. Obwohl physiologische Sensoren bewährte Instrumente zur Klassifizierung von Arbeitsbeanspruchung sind, fehlen ihnen oft kontextbezogene Informationen, um das adaptive System zu informieren. In anderen Studien wurden deshalb Blick- und Interaktionsmessungen in Kombination mit modellbasierten Rückschlüssen auf den Ressourcenbedarf und die mentale Arbeitsbeanspruchung verwendet. Zum Beispiel schätzten Maiwald und Schulte (2012) den mentalen Ressourcenbedarf eines Hubschrauberpiloten auf der Grundlage von Interaktions- und Blickmessungen. Sie passten die Modalität eines Alarmierungssystems an die Verfügbarkeit einer Wahrnehmungsressource an. Wenn die visuelle Belastung hoch war, wurden alle Benachrichtigungen in Form von Sprach- oder Tonmeldungen ausgegeben. Die experimentelle Auswertung in einem Simulator zeigte eine Leistungssteigerung und gute subjektive Bewertungen.

Brand und Schulte (2021) entwickelten ein adaptives Unterstützungssystem für Hubschrauberpiloten, das auf der Messung und Ableitung der aktuellen Aktivität des Piloten nach Honecker und Schulte (2017) basiert. Die Arbeitsbeanspruchung wurde auf der Grundlage der Aktivität des Piloten und eines Aufgabenmodells geschätzt, das den verschiedenen Aufgaben psychische Anforderungen zuordnet. Der Grad der Unterstützung wurde in Abhängigkeit von der Kritikalität der Situation und dem Aufgabenkontext angepasst. In einem Experiment mit Berufspiloten wurde die Unterstützung als angemessen und hilfreich bewertet.

### Leistungsadaptive Systeme

Zwar ist Leistung an sich kein kognitiver Zustand, wird aber im Fall von Leistungseinbrüchen oder Fehlern als Hilfsgröße für einen suboptimalen kognitiven Zustand des Nutzers herangezogen. Hier ist der Vorteil, dass im Gegensatz zur Messung der Arbeitsbeanspruchung, die Leistungsmessung leicht mithilfe von Laboraufgaben umgesetzt werden kann. Kidwell et al. (2012) verglichen adaptive mit einer adaptierbaren Automation in einer Überwachungsaufgabe für mehrere autonome Fahrzeuge. In der adaptierbaren Automatisierungsbedingung konnte der Unterstützungsgrad in einer Bildanalyseaufgabe durch den Bediener angepasst werden. In der adaptiven Automatisierungsbedingung wurde der Unterstützungsgrad im Hinblick auf die Aufgabenerfüllung geändert. Die Autoren stellten fest, dass die adaptive Automatisierung einen leichten Anstieg der Zeit bis zum Abschluss der experimentellen Aufgabe bewirkte, jedoch mit einer geringeren Leistung bei der Änderungserkennung. Außerdem wählten die Bediener während des Versuchs überwiegend einen mittleren Automatisierungsgrad. Calhoun et al. (2011) verwendeten Leistungsmessungen aus mehreren Aufgaben, um eine adaptive Automatisierung mit mittlerem oder hohem Automationsgrad bei der Bedienerunterstützung auszulösen. Dieser Mechanismus wurde mit statischer Automatisierung mit niedrigem Automationsgrad verglichen. Die Ergebnisse zeigten eine subjektive Präferenz für adaptive Automatisierung sowie Leistungsverbesserungen bei mehreren Aufgaben. Tjerk et al. (2010) evaluierten adaptive Automatisierung in einem *Wizard-of-Oz*-Experiment. Die Bedienerunterstützung wurde sowohl an die Leistungsmessung als auch an die anstehenden Aufgaben angepasst. Beim Vergleich zwischen adaptiver Unterstützung und keiner Unterstützung stellten sie eine große Verbesserung der Aufgabenleistung fest. Überraschenderweise war die subjektive Arbeitsbelastung in beiden Bedingungen gleich. Die Autoren erklärten dies damit, dass die professionellen Teilnehmer ihre Strategie anpassten, um ihre Arbeitsbeanspruchung auf einem mittleren Niveau zu halten. In einer adaptiven Automatisierungsbedingung bewerteten Parasuraman et al. (2009) die Leistung des Bedieners bei der Erkennung von Veränderungen, um die automatische Zielerkennung auszulösen, und verglichen sie mit der manuellen Steuerung und der statischen Automatisierung. Sowohl die statische als auch die adaptive Automatisierung steigerten die Sicherheit und Leistung des Bedieners und verringerten gleichzeitig die mentale Arbeitsbeanspruchung. Darüber hinaus waren die positiven Auswirkungen bei der adaptiven Automatisierung sogar noch größer.

### Situationsbewusstseinsadaptive Systeme

In den erwähnten Studien mit beanspruchungs- und leistungsbasierten Auslösern passen die Systeme in der Regel ihren Grad der Ausführungsunterstützung an die Aufgaben an. Handelt es sich bei der Aufgabe beispielsweise um die Klassifizierung eines Sensorbildes, schlagen die adaptiven Systeme mögliche Klassifizierungen vor. Diese Studien gehen davon aus, dass es möglich ist, die Ausführung der Aufgabe zu automatisieren. Im Gegensatz dazu können Auslöser, die auf Messungen der Aufmerksamkeit und des Situationsbewusstseins basieren, auch dann verwendet werden, wenn die Aufgabenausführung nicht automatisiert werden kann oder soll. Die Überlegung hinter diesen Studien

ist, dass sich die Leistung bei der Ausführung verbessert, wenn die SB des Bedieners hoch ist oder die Aufmerksamkeit auf die relevantesten Aufgabeninformationen gerichtet ist.

Zum Beispiel entwickelten Fortmann und Mengerlinghaus (2014) eine auf Blickbewegungsmessung basierende adaptive Schnittstelle zur Unterstützung einer UAV-Überwachungsaufgabe. Sie maßen die Blickbewegungen der Teilnehmer und schlossen daraus, welche Informationen des Systems nicht richtig überwacht werden. Basierend auf dieser Messung wurde die Hervorhebung relevanter Objekte auf der taktischen Karte angepasst. Die experimentelle Auswertung ergab eine Steigerung der Überwachungsleistung bei gleichzeitig höherer Arbeitsbeanspruchung unter den adaptiven Bedingungen. Die Autoren berichteten auch, dass die Bediener das System für nützlich hielten, wobei die Kalibrierung der Erkennbarkeit noch verbessert werden könnte. Bosse et al. (2009) verwendeten ein Aufmerksamkeitsmodell, das Blickmessung und Displaymerkmale integriert, um die Aufmerksamkeit des Benutzers zu schätzen. Gleichzeitig errechnet ein Software-Agent die gewünschte Aufmerksamkeit auf der Grundlage des Wissens über die Aufgabe. In einem Experiment mit einer Aufgabe der Seekriegsführung verwendeten die Autoren diese Methode, um die Hervorhebung von Objekten auf einer Karte auf der Grundlage der Diskrepanz zwischen tatsächlicher und gewünschter Aufmerksamkeit anzupassen. Die Autoren stellten eine signifikante Verbesserung fest, als sie keine Unterstützung mit der adaptiven Unterstützung verglichen.

Im Zusammenhang mit Flugzeugcockpits entwickelten Lounis, Peysakhovich und Causse (2020a) ein Unterstützungssystem auf der Grundlage von Blickbewegungsmessung. Durch den Vergleich der Verweildauer der Piloten auf verschiedenen Cockpitanzeigen mit einer Datenbank des Standard-Blickverhaltens erkannte das adaptive System eine von der Norm abweichendes Überwachungsverhalten und löste einen akustischen Alarm aus. Obwohl das System in der Lage war, die Aufmerksamkeit auf kritische Fluginstrumente umzulenken, wurden keine Leistungsverbesserungen festgestellt, und die subjektive Bewertung war aufgrund von Fehlalarmen negativ. Die Autoren kamen zu dem Schluss, dass die Integration von Flugparametern die Benutzerfreundlichkeit des Systems verbessern könnte.

## Hybride Systeme

Andere Studien verwendeten eine Kombination verschiedener Maßnahmen, um adaptive Systeme zu implementieren. Schwarz und Fuchs (2017) quantifizierten den Zustand des Nutzers in den zwei Dimensionen Beanspruchung sowie Aufmerksamkeit und verwendeten eine Kombination aus physiologischen, verhaltensbezogenen und leistungsbezogenen Messungen, um kritische Zustände zu identifizieren. In einer Studie evaluierten sie fünf Adaptionsstrategien für verschiedene Aufgaben in einer Marine-Luftüberwachungsaufgabe. Beim Vergleich der Strategien mit einer Kontrollperiode stellten sie fest, dass ihre Interventionen die Gesamtzahl der kritischen Zustände nicht verringerten, aber die Dauer des kognitiven Konflikts reduzierten, der die Intervention auslöste. Sie berichteten auch, dass die Wirksamkeit der einzelnen Anpassungen möglicherweise auf ihren spezifischen Anwendungsfall beschränkt sind und andere Aufgaben beeinträchtigen könnte.

### 2.4.3 RESÜMEE

Der erste Teil dieses Abschnitts definiert zunächst, was kooperative Automation ist, und wie diese bei der Reduzierung menschlicher Fehler helfen könnte. Dabei wird hergeleitet, dass Adaptivität im Kern der Kooperation von Menschen und Automation liegt. Ein adaptives System besteht aus einer Strategie, die wiederum aus einem Auslösemechanismus und einer Anpassung des Systemverhaltens zusammengesetzt ist.

Unabhängig von der verwendeten Strategie gibt es in der Literatur zwei verschiedene Ansätze für die Anpassung des Verhaltens. Beim ersten Ansatz sind Auslösemechanismus und Adaption unabhängig voneinander wie zum Beispiel die Messung der Herzfrequenz zur Erhöhung des Automationsgrads einer Systemfunktion (vgl. Aricò et al. (2016) und Wilson und Russell (2007)). Die Adaption selbst ist inhaltlich unabhängig vom in der Adaptionstrategie verwendeten Maß. Beim zweiten Ansatz wird die Information des Auslösemechanismus verwendet, um eine kontextbezogene Adaption umzusetzen. Ein Beispiel hierfür wäre ein System, das feststellt, ob ein Bediener gerade an einer Aufgabe arbeitet und dann basierend darauf, andere Aufgaben übernimmt (vgl. Brand und Schulte (2021)). Hierbei muss das System bei der Adaption also Kontextwissen verarbeiten, welches aus dem Auslösemechanismus stammt. Die empirischen Ergebnisse der diskutierten Studien zeigen, dass der zweite Ansatz in Hinsicht auf die Akzeptanz vielversprechender ist, da der Mehrwert direkt von den Bedienern anerkannt wird.

Der Überblick über die Studien zeigt, dass adaptive System basierend auf der Erkennung des kognitiven Zustands eines Bedieners ein sinnvoller Ansatz zur Reduktion des menschlichen Fehlers ist. Die kognitiven Zustände, auf deren Basis eine Anpassung getroffen wird, basieren meist auf Modellen von Arbeitsbeanspruchung, der damit korrelierten Leistung oder SB. Für das Konstrukt der mentalen Beanspruchung haben Brand und Schulte (2021) zeigen können, dass eine kontextreiche Operationalisierung sinnvoll für den Einsatz in einem adaptiven Assistenzsystem ist. Die Art der Adaption ist bei Beanspruchungsmessung oft eine Anpassung der Automationsgrade, während bei Systemen auf Basis von SB-Messungen die Anzeige adaptiert beziehungsweise der Nutzer benachrichtigt wird.

Die Studien im Zusammenhang mit SB zeigen, dass eine SB-adaptive Anpassung der Anzeige vielversprechend ist für Aufgaben im Bereich der Überwachung und Bemerkens aufgabenrelevanter Informationen. Jedoch sind die Modelle und die Implementierung der Interventionen auf die jeweiligen Anwendungsfälle zugeschnitten, was die Übertragung auf andere Arbeitssituationen erschwert. Hier fehlt es an einem grundlegenden Konzept für eine maschinell messbaren Operationalisierung von SB sowie einer generischen Interventionsstrategie. Aus diesen Lücken leiten sich die zwei zentralen Forschungsfragestellungen für die vorliegende Arbeit ab:

- Wie kann Situationsbewusstsein operationalisiert und maschinell gemessen werden, sodass es in einem adaptiven System verwendet werden kann?
- Wie sollte ein adaptives Assistenzsystem, das auf das Situationsbewusstsein des Bedieners reagiert, gestaltet sein, um die Leistung des Mensch-Maschine-Systems zu verbessern?

Im folgenden Kapiteln werden diese Forschungsfragen adressiert. Das nächste Kapitel beschreibt das technische Konzept und im Anschluss werden die durchgeführten Validierungsstudien beschrieben.

### 3 DAS SITUATIONSBEWUSSTSEINSADAPTIVE ASSISTENZSYSTEM

Dieses Kapitel beschreibt das Konzept eines situationsbewusstseinsadaptiven Assistenzsystems. Als Zielvorstellung dient das Beispiel aus der Einleitung (siehe Kapitel 1) vom Beifahrer, der den Autofahrer warnt, wenn relevante Informationen nicht bemerkt wurden. Im Folgenden wird zunächst das adaptive Systemverhalten des Systems genauer spezifiziert, welches den Ausgangspunkt für die Konzeption des Systems darstellt. Anschließend werden die zur Umsetzung notwendigen Funktionen in den Abschnitten 3.2 - 3.4 beschrieben. Die Architektur des Gesamtsystems wird dann abschließend in Abschnitt 3.5 beschrieben. Der Inhalt dieses Kapitels stehen in Zusammenhang mit mehreren veröffentlichten Studien (Schwerd & Schulte, 2020, 2021a, 2021b, 2021c).

#### 3.1 ADAPTIVES SYSTEMVERHALTEN

Ziel des Assistenzsystems ist es, Leistungsdekremente (wie z.B. Fehlentscheidungen), die auf niedriges Situationsbewusstsein zurückzuführen sind, mit einer adäquaten Intervention zu vermeiden. Das Assistenzsystem soll für Arbeitsplätze zur Steuerung dynamischer, technischer Systeme geeignet sein wie beispielweise ein Auto, Flugzeug oder Anlagenprozess. In diesen Arbeitsumgebungen muss ein Bediener oft mehrere Aufgaben gleichzeitig verfolgen unter Beachtung verschiedener Informationsquellen. In solchen Arbeitsumgebungen wird das gewünschte Systemverhalten des Assistenzsystems wie folgt definiert:

- 1) Das Assistenzsystem soll den Bediener auf Informationen hinweisen, die er **nicht wahrgenommen hat und ihm nicht bewusst** sind, und damit das Situationsbewusstsein und die Leistung des Gesamtsystems verbessern.
- 2) Dabei soll das System sowohl auf die Existenz **neuer Informationen** hinweisen als auch auf **Veränderungen von Systemzuständen**, die nicht vom Bediener wahrgenommen wurden.

Aus der gewünschten Arbeitsumgebung sowie den beiden Verhaltensgrundsätzen können folgende spezifischen Anforderungen an das Assistenzsystem abgeleitet werden:

- Das System muss kontinuierlich **die Aufmerksamkeit** des Bedieners abschätzen und darauf aufbauend das **Situationsbewusstsein** des Bedieners bewerten.
- Die Bewertung des SBs muss so gestaltet sein, dass eine **informationsspezifische Auskunft** darüber möglich ist, welche Systeminformationen im SB fehlen oder falsch sind. Das System muss dabei **zwischen Unwissen und falschem Wissen unterscheiden** können.
- Das System muss **relevante Informationen von nicht relevanten Informationen** unterscheiden, um auch in Mehrfachaufgabenumgebungen unterstützen zu können.
- Das System muss dem Bediener über eine geeignete Schnittstelle **benachrichtigen und auf Informationen aufmerksam machen** können.

Zur Erfüllung dieser Anforderungen wird ein Systemaufbau vorgeschlagen, der sich an den vier Funktionen eines informationsverarbeitenden Systems nach Parasuraman et al. (2000) orientiert. Ein solches System hat die vier Komponenten Informationserwerb, Informationsbewertung, Entscheidungsfindung und Aktionsausführung. Übertragen auf das vorliegende Assistenzsystemkonzept ist die Funktion des Informationserwerbs die **Aufmerksamkeitsanalyse**, das

durch kontinuierliche Verhaltensbeobachtung des Bedieners eine Abschätzung darüber trifft, auf welchen Informationen die Aufmerksamkeit des Bedieners gerade liegt. Die Ergebnisse dieser Messung verarbeitet das zweite Modul **Situationsbewusstseinsbewertung** (vgl. Informationsanalyse) und erstellt ein dynamisches Modell, welches Aussagen über das SB des Bedieners in Bezug auf Arbeitskontext sowie den tatsächlichen Situationsstand treffen kann. Das Ergebnis dieser Bewertung erhält das Entscheidungsfindungs-Modul **Interventionsgenerierung**, welche den Bediener im Kontext der gegenwärtig wichtigsten Aufgaben auf ihm nicht bewusste Informationen hinweist. Die Umsetzung der Intervention (vgl. Aktionsausführung) obliegt in diesem Fall der Mensch-Maschine-Schnittstelle, deren Umsetzung die konventionelle Funktion einer MMS darstellt und im Folgenden nicht detailliert beschrieben wird. Abbildung 3-1 zeigt die Module zusammen mit den folgenden Unterkapitelnummern, in denen die einzelnen Module detailliert beschrieben werden. Im Anschluss an die Beschreibung der Module, zeigt der Abschnitt 3.5 den Gesamtsystemaufbau.

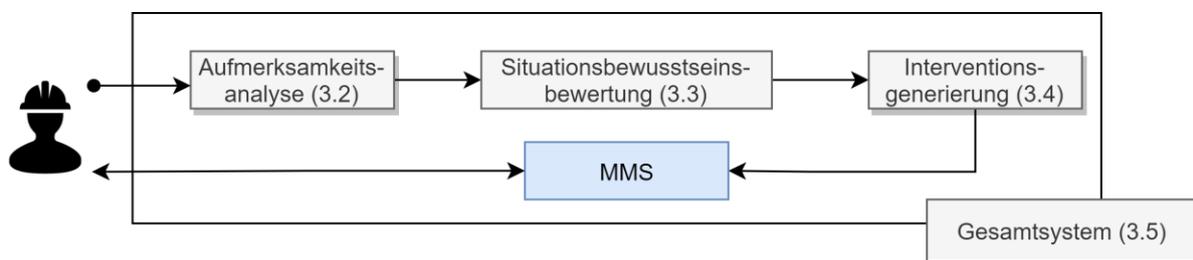


Abbildung 3-1 - Struktur des adaptiven Systems bestehend aus Aufmerksamkeitsanalyse, SB-Bewertung und Intervention, die sich auf die Mensch-Maschine-Schnittstelle (MMS) auswirkt.

### 3.2 AUFMERKSAMKEITSANALYSE

Eine Aufmerksamkeitsanalyse in einer dynamischen Arbeitssituation muss möglichst robust und in Fast-Echtzeit bestimmen können, welche Informationen vom Bediener vollständig verarbeitet wurden. Das bedeutet, dass der Bediener eine Information wahrgenommen sowie verarbeitet hat und diese ihm unmittelbar zur Entscheidungsfindung bei einer Aufgabe zur Verfügung stehen. Wie im Abschnitt 2.1.4 ausgeführt, ist die Blickbewegungsmessung eine bewährte Methode zur Aufmerksamkeitsanalyse, welche für die Umsetzung des hier beschriebenen Systems als Messverfahren dienen soll.

Die konventionellen Methoden zur Auswertung von Blickbewegungsdaten, wie beispielsweise die Analyse der Fixationsfrequenz auf vordefinierte Interessensbereiche liefern in der Arbeitsumgebung keine Informationen zur Verwendung in einem Assistenzsystem. Es nützt beispielsweise nichts, eine optimale durchschnittliche Verteilung der Blickrichtung des Bedieners zu wissen, wenn die Aufmerksamkeit in einer kritischen Situation an der falschen Stelle liegt. Aus diesem Grund wird Aufmerksamkeitsanalyse im vorliegenden Konzept als **Klassifikationsproblem** betrachtet mit zwei zentralen Fragen:

- Wurde eine Information wahrgenommen, die innerhalb eines vom Bediener fixierten Interessensgebiets liegt?
- Welcher Informationsstand wurde wahrgenommen?

Abbildung 3-2 zeigt die Messkette zur kontinuierlichen Aufmerksamkeitsanalyse. Eingangsgrößen der Messkette sind die Blickbewegungsmessung und Interaktionsmessung.

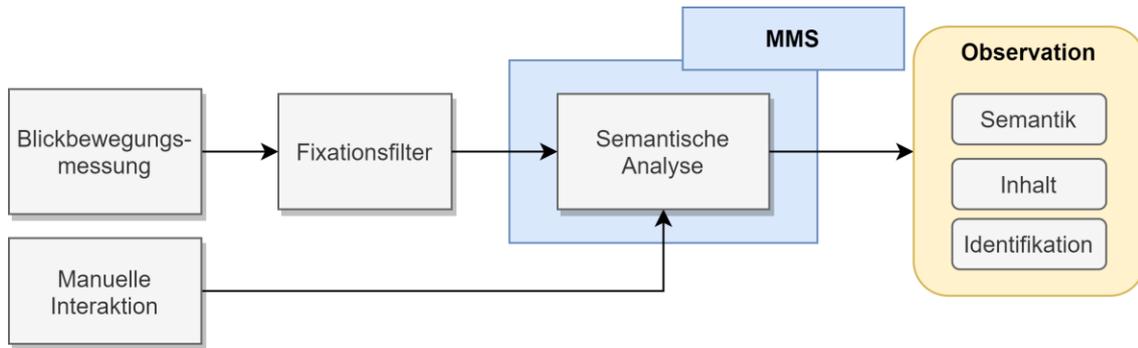


Abbildung 3-2 - Messkette zur kontinuierlichen Analyse von visueller Aufmerksamkeit.

Die gewünschte Ausgangsgröße der Aufmerksamkeitsanalyse ist die vom Bediener wahrgenommenen Informationen. Diese Datenpunkte werden **OBSERVATIONEN** genannt und bestehen aus drei Variablen:

- **SEMANTIK:** Die Semantik beschreibt das fixierte **Interessensgebiet** mit einem eindeutigen Namen. Es ist somit die Bedeutung dessen, was wahrgenommen wurde, wie beispielsweise eine „Geschwindigkeitsanzeige“.
- **INHALT:** Der Inhalt beschreibt welcher **Typ** und **Stand** der im Interessensgebiet dargestellten Information zum Zeitpunkt der Fixation dargestellt war, wie beispielsweise „120 km/h“.
- **IDENTIFIKATION:** Die Identifikation beschreibt, **welchem Objekt** die Information zugeordnet ist. Diese Variable ist besonders nützlich, wenn in einer Aufgabenumgebung Informationen gleicher Semantik für verschiedene Objekte vorhanden sind, wie beispielsweise die Geschwindigkeit anderer Autos im Straßenverkehr.

Zur Generierung von **OBSERVATIONEN** werden sowohl Blickbewegungsmessung als auch Interaktion verarbeitet, was im Folgenden beschrieben wird.

### 3.2.1 VERARBEITUNGSKETTE DER BLICKBEWEGUNGSMESSUNG

Die Datenverarbeitung der Blickbewegungsmessung basiert auf einem Fixationsfilter, der anschließenden Verarbeitung der Fixationen und der semantischen Analyse der Messpunkte in einem gegebenen Auswertungsradius.

#### Fixationsfilter

Da ein Bediener Informationen vorwiegend mithilfe von Fixationen wahrnimmt (siehe 2.1.4.3), werden aus allen gemessenen Blickpunkten nur Fixationen weiterverarbeitet, während Sakkaden sowie Blinzeln herausgefiltert werden. Kommerzielle Blickbewegungsmesssysteme bieten häufig Klassifikationsmethoden an, mit denen ein Blickbewegungspunkt der Kategorie Fixation, Sakkaden oder Blinzeln zugeordnet wird. Ferner werden nur Fixationen weiterverarbeitet, die eine Anzeigeoberfläche schneiden und somit den dort angezeigten Informationen zugeordnet werden können.

#### Verarbeitungsradius

Abgeleitet aus den Erkenntnissen über periphere Wahrnehmung (siehe 2.1.4.3), wird nicht nur die Fixationsposition weiterverarbeitet, sondern die Peripherie mit ausgewertet. Diesen Ansatz haben

bereits Honecker und Schulte (2017) umgesetzt und gezeigt, dass die nah-periphere Auswertung des fixierten Blickpunkts Vorteile für die Genauigkeit der Aufmerksamkeitsanalyse haben kann. Hierfür wird ein kreisförmiger Bereich um die Fixation ausgewertet, der zur Klassifikation von Aufmerksamkeit auf ein Interessensgebiet um eine Fixation herum angewendet wird. Abbildung 3-3 macht den Unterschied zwischen zwei Radien deutlich, wenn eine Fixation in der Nähe eines Interessensgebiets liegt. In der Abbildung schneidet nur Radius 2 das Interessensgebiet (*Area of Interest* – Aol 1) und generiert somit eine **OBSERVATION**. Bei der Zuweisung einer Fixation zu einem Interessensgebiet bei gegebenen Auswertungsradius wird die euklidische Distanz zur nächsten Kante des Interessensgebiets berechnet.

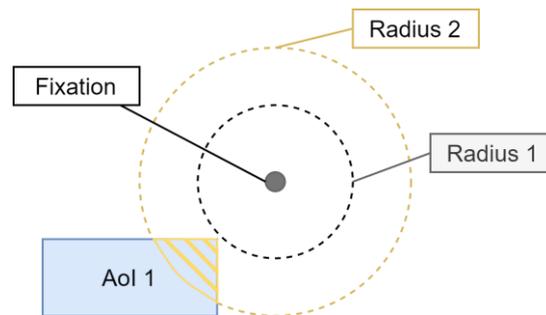


Abbildung 3-3 - Vergleich einer Fixation mit zwei Auswertungsradien. Beim Radius 1 liegt die Aol 1 nicht innerhalb des Auswertungsradius und wird somit nicht dieser Fixation zugewiesen. Bei Radius 2 liegt die Aol 1 innerhalb des Radius und wird somit als Fixation auf Aol 1 verarbeitet.

Dies berücksichtigt zum einen das hochauflösende Sichtfeld des Menschen und kann zum anderen auch Messungenauigkeiten ausgleichen. Letzteres gilt in vielen Fällen, bei denen die gemessene Fixation wenige Grad neben einem Interessensgebiet im „informationsleeren“ Raum liegt. Dass Fixationen in Arbeitssituationen meist nicht ins Leere, sondern auf Gebieten mit Informationen liegen, wurde beispielsweise von Hornof und Halverson (2002) genutzt um die Genauigkeit von Blickbewegungsdaten zu überwachen und zu bewerten. Nachteil dieses Ansatzes ist, dass hierbei auch mehrere Aols innerhalb des Radius liegen können. An dieser Stelle muss die Entscheidung getroffen werden, ob das räumlich am nächsten liegende Interessensgebiet verarbeitet wird oder alle im Radius liegenden Gebiete weiterverarbeitet werden. In der vorliegenden Arbeit wurde der letztere Ansatz gewählt.

### Semantische Analyse

Die Fixationen werden weitergegeben an die semantische Analyse, in der ausgewertet wird, was an der fixierten Position dargestellt ist. Ein Beispiel zeigt die Abbildung 3-4, in der zwei hypothetische Fixationen in einem Autocockpit dargestellt sind. Die erste Fixation liegt auf dem Tacho, auf dem 108 km/h Knoten dargestellt sind. Die resultierende **OBSERVATION** besteht hier aus der **SEMANTIK** „Geschwindigkeit“ und dem **INHALT** 30 m/s. Die Umrechnung von km/h in m/s dient der einheitlichen Darstellung von Geschwindigkeiten zur Weiterverarbeitung. Diese **OBSERVATION** ist mit dem eigenen Auto des Fahrers assoziiert, wofür hier die **IDENTIFIKATION** mit dem generischen Namen „Auto“ steht.

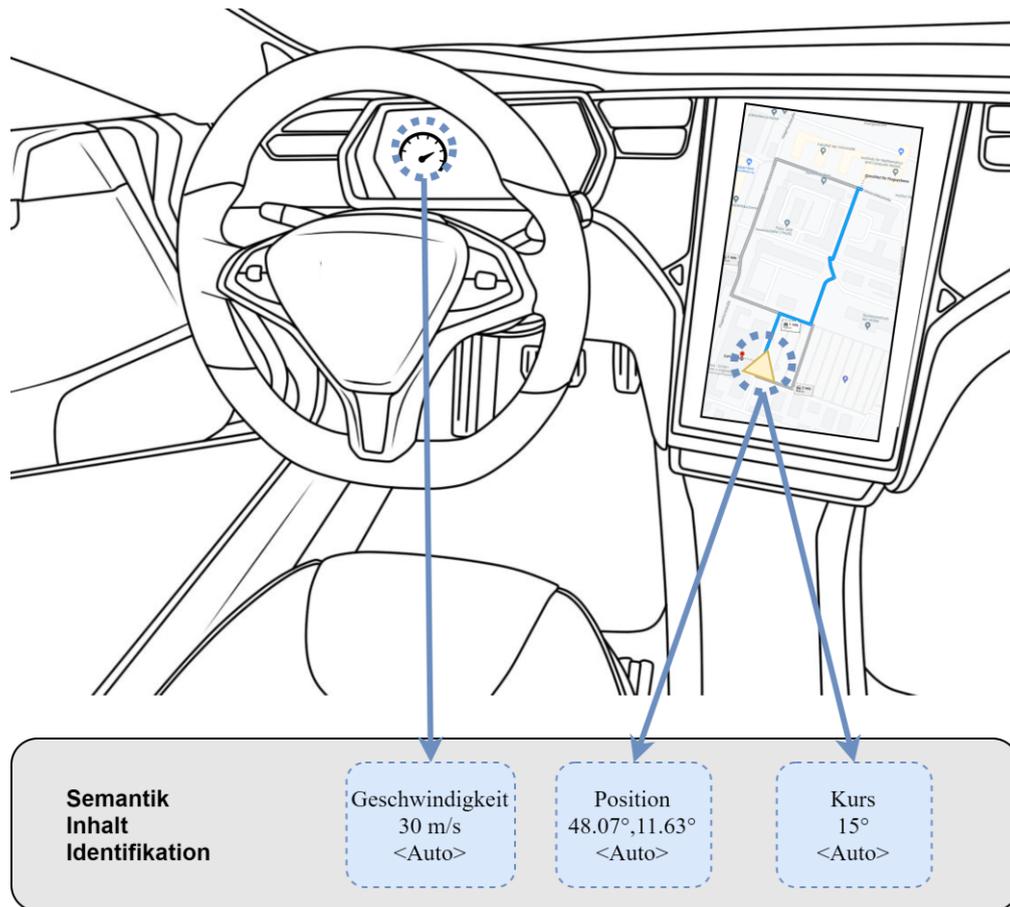


Abbildung 3-4 - Beispiel der semantischen Auswertung zweier hypothetischer Fixationen im Cockpit eines Autos.

Bei der Fixation auf die Navigationskarte zeigt sich ein weiterer Aspekt, der bei der semantischen Analyse bedacht werden muss. Es gibt Anzeigeelemente, welche mehr als nur eine Information an einer Stelle enkodieren. Beispielsweise ist beim Blick auf das Navigationsdisplay sowohl die Position des Autos relativ zur Umgebung dargestellt als auch die Orientierung des fahrenden Autos. Dies resultiert in zwei **OBSERVATIONEN** mit der **SEMANTIK** „Position“ und „Kurs“, der **IDENTIFIKATION** des Autos sowie den dargestellten **INHALTEN**.

### 3.2.2 AUFMERKSAMKEITSBESTIMMUNG DURCH MANUELLE INTERAKTION

Die Beobachtung der manuellen Interaktion ist in softwarebasierten Bediensystemen eine einfache umzusetzende Variante zur Aufmerksamkeitsanalyse. Sie ist aber nicht ausreichend, wenn die Bedienung des Systems viele Aufgaben aus dem Bereich der überwachenden Kontrolle erfordert und damit nicht kontinuierlich manuelle Interaktionen zu erfolgen haben. Hier kann die manuelle Interaktion als Ergänzung zum Blickbewegungssystem gesehen werden und dient vor allem der Robustheit des Systems, da es bei der Verwendung von gängigen Blickbewegungsmesssystemen durchaus vorkommen kann, dass die Blickrichtung zeitweise nicht erkannt wird oder in bestimmten Bereichen des Arbeitsplatzes ungenau ist (Hornof & Halverson, 2002).

Die manuelle Bedienung des Systems ist ein starker Indikator für visuelle Aufmerksamkeit auf bestimmte Orte im Arbeitsplatz. Diese Annahme hat sich in Zusammenhang mit zahlreichen Arbeiten zur Blickbewegungskalibrierung als valide herausgestellt. Hierbei wurde einerseits das Klicken mit einer Computermaus (M. X. Huang et al., 2016; Sugano et al., 2008) und andererseits die Bedienung von berührungssensitiven Bildschirmen untersucht (X. Zhang et al., 2018). Bei diesen beiden

Anwendungsfällen ist der Zusammenhang besonders ausgeprägt, da Positionierung der Klick- oder Berührungspunkte visuelle Aufmerksamkeit auf Mausymbol oder Finger notwendig ist. Bei haptischen Bedienelemente wie beispielsweise Schalthebel oder Gaspedal ist dieser Zusammenhang nicht gegeben. Mit Blick auf die Anwendung im Cockpit kann, kann also die Interaktionsbeobachtung bei berührungssensitiven Bildschirmen verwendet werden, um die Messung der visuellen Aufmerksamkeit auf Basis der Blickbewegungsmessung zu ergänzen.

Die semantische Analyse der Interaktionspunkte ist äquivalent zur Verarbeitung von Fixationspunkte und generiert zum Zeitpunkt der Interaktion eine **OBSERVATION** mit den drei bereits beschriebenen Datenpunkten.

### 3.3 SITUATIONSBEWUSSTSEINSBEWERTUNG

Die Aufmerksamkeitsanalyse generiert mit einem Strom an **OBSERVATIONEN**, welche fixierte Informationen und deren Inhalte repräsentieren. Der folgende Abschnitt diskutiert, wie aus diesem Messdatenstrom eine Bewertung des Situationsbewusstseins getroffen werden kann.

Hierfür muss zunächst eine Operationalisierung von SB gefunden werden, die maschinell repräsentierbar ist und mithilfe der Messdaten generiert werden kann. Als Grundlage hierfür dient das Modell von Endsley (siehe Abschnitt 2.2.1.1), in dem domänenspezifische Systeminformationen genutzt werden, um das Situationsbewusstsein zu beschreiben. Es ist also vorteilhaft, wenn das Modell Kontextinformationen über das System beinhalten kann. Darüber hinaus wird Situationsbewusstsein als normatives Modell angenommen, das heißt, dass es eine Repräsentation des Modells gibt, die als korrekt definiert ist. Wenn beispielsweise ein Bediener in einem Versuch nach einer Systemgröße (bspw. Geschwindigkeit des Autos) gefragt wird, kann die Antwort mit der tatsächlichen Geschwindigkeit verglichen werden, um das Maß der Richtigkeit zu quantifizieren. Diese Vorstellung ermöglicht die Quantifizierung von SB und ist damit ein guter Ausgangspunkt für ein Messsystem. Es gilt aber zu beachten, dass diese normative Perspektive des Modells durch einige Wissenschaftlicher als wenig hilfreich kritisiert wurde (Stanton et al., 2017).

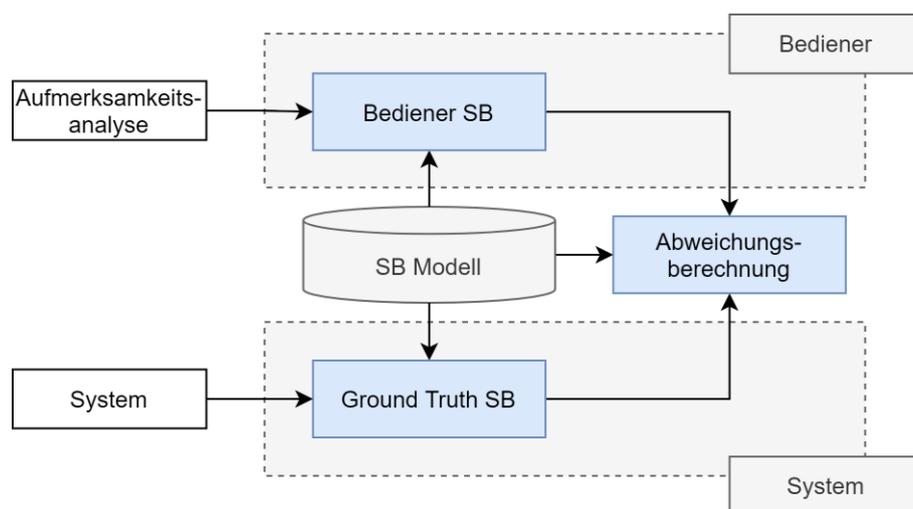


Abbildung 3-5 - Messkette zur Bewertung von Situationsbewusstsein.

Die Abbildung 3-5 zeigt die Messkette, deren zwei Eingangsgrößen das Ergebnis der Aufmerksamkeitsanalyse ist und der aktuelle Systemzustand. Die zugrundeliegende Idee ist, dass die Analyse der Bedieneraufmerksamkeit genutzt wird, um festzustellen, welcher Stand des Systems dem Bediener bewusst ist. Gleichzeitig wird eine analoge Repräsentation eines „perfekten“

Systemüberwachers erstellt, der alle Systemzustände gleichzeitig beobachten kann. Diese beiden Repräsentationen werden dann verglichen, um fehlendes oder falsches Situationsbewusstsein beim Bediener feststellen zu können. Die einzelnen Module *SB Modell*, *Bediener SB*, *Ground Truth SB* und *Abweichungsberechnung* werden im Folgenden genauer beschrieben.

### 3.3.1 SITUATIONSBEWUSSTSEINSMODELL

Das im Folgenden beschriebene Modell ist eine Kombination der Arbeiten von R. E. T. Jones et al. (2011) und Frische und Lüdtkke (2013). Erstere formulierten ihr Modell mithilfe einer *Fuzzy Cognitive Map* (FCM), in dem ein Knoten jeweils einer spezifischen Information der Domäne zugeordnet war oder einen aggregierten Informationsstand repräsentierte (siehe Abschnitt 2.2.1.3). Die Beziehungen zwischen Informationen sowie Assoziationen mit Arbeitszielen wurden über gewichtete Kanten zwischen den Knoten dargestellt. Die Gewichtungen wurden dann in einem Inferenzschritt genutzt, um ausgehend von Eingangsgrößen Aussagen über das geschätzte Situationsbewusstsein zu treffen, die in dieser Studie simuliert wurden. Frische und Lüdtkke (2013) erstellten ein normatives Situationsmodell, welches alle vorhandenen Informationen im System dynamisch repräsentiert und verglichen dieses mit einer Bewertung des SBs des Bedieners.

Für das hier beschriebene Modell wird die Formulierung mittels eines Netzwerks aus Knoten und Kanten übernommen ohne die Gewichtung der Verbindungen. Der Grund hierfür ist, dass die Wahl dieser Gewichte mit viel Unsicherheit behaftet ist und eine falsche Wahl in der komplexen Systemdynamik des FCM schwer zu validieren ist. In einer Studie mit FCMs, die im Kontext dieser Arbeit angefertigt wurde, wurden deshalb zunächst nur wenige diskrete Gewichtung gewählt um diese Komplexität zu beherrschen (Schwerd & Schulte, 2021b). Diese Studie zeigte aber auch, dass Gewichtungen allein nicht die gewünschte Struktur für das Situationsbewusstseinsmodell abbilden können. Ein Situationsbewusstseinsmodell muss in der Lage sein, objektspezifische Daten parallel abzubilden und gegebenenfalls auch in Zusammenhang zu setzen. Ein konkretes Problem kommt hier aus der Anwendungsdomäne und dem Informationsraum, in dem man sich in militärischen Führungssystemen zurechtfinden muss. Hier muss ein Bediener Informationen der gleichen Semantik (z.B. Position) verschiedener Plattformen kennen. Im Kontext einer taktischen Karte bei der Multi-UAV-Führung, wie in Abbildung 3-6 dargestellt, werden beispielsweise die Informationen für die „Position“ zwei verschiedener UAVs dargestellt. Die Distanz zwischen beiden Positionen könnte hier als eine Kombination des Wissens der Position der einzelnen Flugobjekte <UAV1> und <UAV2> modelliert werden. Um dies zu ermöglichen, wird die Beziehung zwischen Informationen und Zielen nicht mittels gewichteter Kanten, sondern durch logische Verbindungen modelliert. Dies entspricht eher dem Ansatz von Frische und Lüdtkke (2013), die ihr Situationsmodell mithilfe logischer Objekte und Regeln formulieren, in der Situationselemente mit eigenen Identitäten verwendet und in Beziehung zueinander gesetzt werden. Aus dieser Studie wurde ebenfalls das Konzept zur Berechnung einer Abweichung zwischen System und Bediener übernommen und die Abweichungsberechnung um ein generisches Modell für unterschiedliche Einheiten erweitert.

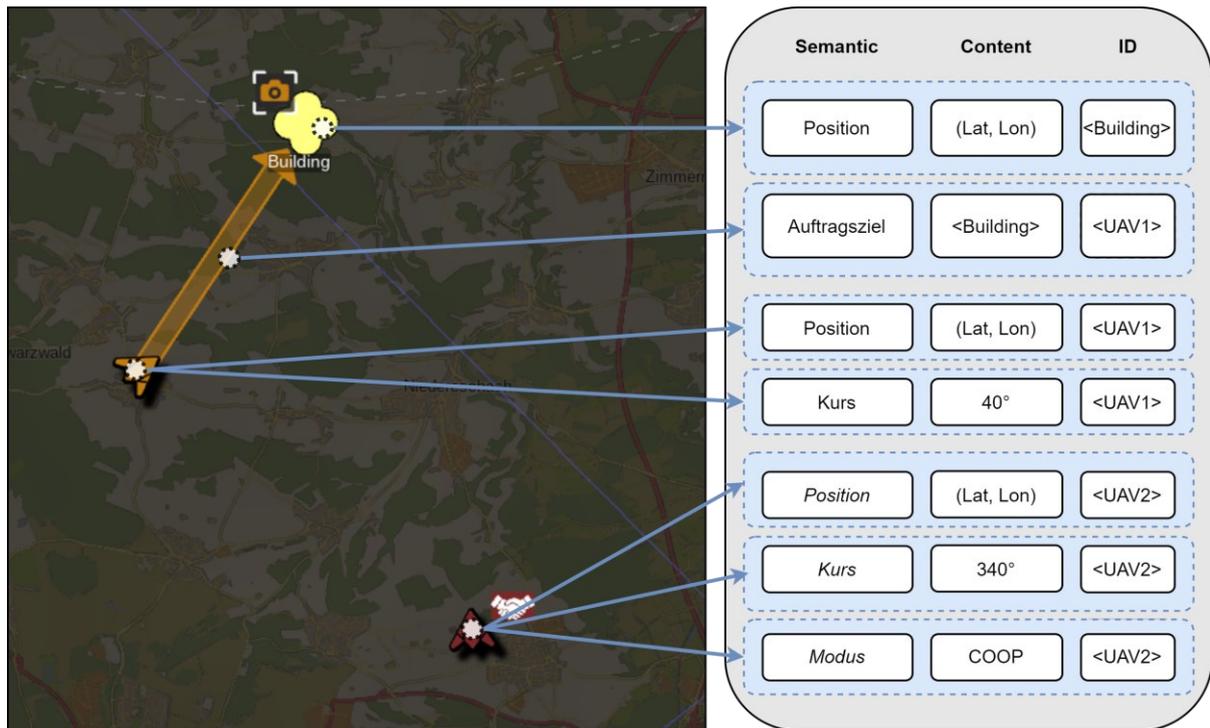


Abbildung 3-6 - Ein Blick auf die taktische Karte mit zwei UAVs und der resultierenden Informationen. Abbildung adaptiert aus Schwerd und Schulte (2021a)

Aufbauend auf diesen Überlegungen soll nun das in dieser Arbeit entwickelte formale Modell beschrieben werden. Bevor das Modell als dynamische Bewertung eingesetzt werden kann, muss eine statische Wissensbasis definiert werden, welches als **SB-MODELL** in Abbildung 3-5 bezeichnet wird. Das **SB-MODELL** besteht aus einem gerichteten Graphen, dessen Knoten für aufgabenrelevante Informationen stehen und dessen Kanten die **RELATIONEN** zwischen diesen Informationen repräsentieren. Die Knoten werden im folgenden **SB-KNOTEN** genannt und stellen entweder spezifische Informationen dar oder stehen für Zusammenhänge oder Aggregationen spezifischer Informationen. Wenn sie spezifische Informationen darstellen, werden die Knoten mit der **SEMANTIK** einer **OBSERVATION** verknüpft, welche diese Information auf der Bedienanzeige darstellt.

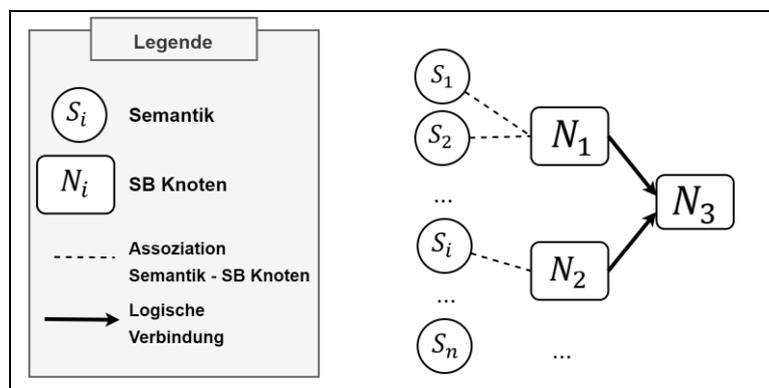


Abbildung 3-7 - Grundlegende Bestandteile des statischen Modells

Für die logischen Verbindungen gibt es drei verschiedene **RELATIONEN**: AND-, OR- oder DIFF-Relationen. Die unterschiedlichen Kantenarten wirken sich auf Inferenz in Abhängigkeit von der **IDENTITÄT** der **OBSERVATIONEN** aus, welche im folgenden Abschnitt beschrieben wird.

### Inferenz des Bediener SB

Das logische Netzwerk dient als Wissensbasis, um basierend auf gemessenen **OBSERVATIONEN** eine dynamische Instanz des SBs des Bediener zu erstellen. Die Inferenz besteht dabei aus drei Schritten:

#### 1. Aktualisierung der Menge von Observationen

Jede erzeugte **OBSERVATION** wird in einer Menge gespeichert, wobei neue **OBSERVATIONEN** vergangene mit der gleichen **SEMANTIK** und **IDENTITÄT** ersetzen.

#### 2. Erstellung und Aktualisierung der mit Observationen assoziierten Knoten

In regelmäßigen Abständen wird für jeden Knoten des **SB-MODELLS** geprüft, ob es eine assoziierte **OBSERVATION** in der Menge der **OBSERVATIONEN** gibt. Wenn ja, wird eine Instanz des Knotens erzeugt und somit „aktiv“ geschaltet. Falls bereits ein **SB-KNOTEN** aktiv ist, wird er mit der neuen **OBSERVATION** aktualisiert. **OBSERVATIONEN** mit der gleichen **SEMANTIK** aber unterschiedlichen **IDENTIFIKATIONEN** erzeugen für jede **IDENTIFIKATION** eine Instanz eines **SB-KNOTENS**. Der **SB-KNOTEN** enthält die aktuelle **OBSERVATION** einschließlich des **INHALTS**, der zum Zeitpunkt der gemessenen Fixation dargestellt war.

#### 3. Weitergabe und Aktualisierung im logischen Netzwerk

Nun wird geprüft, welche **RELATION** alle aktiven **SB-KNOTEN** untereinander haben und nach der folgenden Systematik weitere **SB-KNOTEN** im Netzwerk erstellt oder aktualisiert:

- **AND-RELATION**: Alle Elternknoten müssen existieren und die gleiche **IDENTIFIKATION** vorweisen.
- **OR-RELATION**: Einer der Elternknoten muss existieren. Alle Elternknoten haben die gleiche **IDENTIFIKATION**.
- **DIFF-RELATION**: Alle Elternknoten müssen existieren und unterschiedliche **IDENTIFIKATIONEN** besitzen. Einen Sonderfall gibt es bei der Existenz eines einzigen Elternknoten. Hier wird angenommen, dass es mindestens zwei Elternknoten der gleichen **SEMANTIK** aber mit unterschiedlicher **IDENTIFIKATION** gibt.

Diese aggregierten Knoten erhalten alle **OBSERVATIONEN** ihrer Elternknoten und speichern diese als Menge von assoziierten **OBSERVATIONEN** ab. Dies ist notwendig für die folgende Berechnung von Abweichungen zwischen Systemzustand und SB des Bediener.

Das folgende Beispiel aus der Anwendungsdomäne der Flugverkehrskontrolle illustriert den Inferenzmechanismus. In Abbildung 3-8 ist eine schematische Mensch-Maschine-Schnittstelle (MMS) für einen Fluglotsen mit zwei Flugzeugen (AC1 und AC2) sowie annotierter Flughöhe und -geschwindigkeit dargestellt. Diese Displayelemente sind Interessensgebiete und generieren bei Fixation die assoziierten vier **OBSERVATIONEN**. Beispielsweise wird beim Blick auf das Flugzeugsymbol sowohl Position als auch Kurs als Observation mit der entsprechenden **SEMANTIK** „AC\_Position“ und „AC\_Heading“ generiert.

Das **SB-MODELL** besteht insgesamt aus sieben **SB-KNOTEN**. Die erste Ebene der Knoten sind die Flughöhe, Kurs, Position des Flugzeugs sowie die Position einer Flugverbotszone. Die beiden **SB-KNOTEN**, die den Kurs und die Position eines Flugzeugs betreffen werden über eine **AND-RELATION** zur „Projektion der Flugzeugposition“ zusammengeführt. Dieser Knoten steht dafür, dass der Bediener sofern er Kurs und Position kennt auch die zukünftige Position des Flugzeugs abschätzen. Die Flughöhe ist über eine **DIFF-Relation** mit der Information „Höhentrennung“ verbunden. Das bedeutet, dass zwei **SB-KNOTEN** mit **OBSERVATIONEN** unterschiedlicher **IDENTIFIKATIONEN** den Knoten

„Höhentrennung“ erstellen, also die Kenntnis über die Höhe zweier Flugzeuge auch eine Kenntnis über deren vertikale Trennung bedeutet. Außerdem sind die **SB-KNOTEN** Flugzeugposition und Position Flugverbotszone über eine **DIFF-RELATION** verbunden, was modelliert, dass das Wissen über diese beiden Informationen die Distanz zwischen beiden suggeriert.

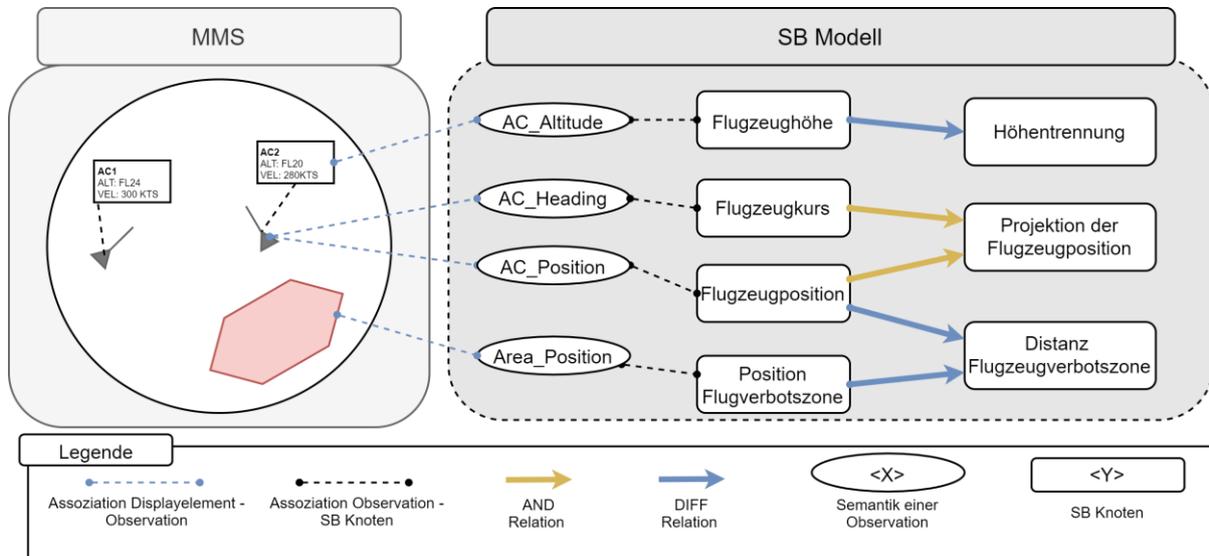


Abbildung 3-8 - Beispiel eines SB-Modells mit Knoten und den assoziierten Observationssemantiken.

Aufbauend auf diesem Modell zeigt Abbildung 3-9 beispielhaft, wie die Inferenz des Bedienersituationsbewusstseins aussieht. Es wird angenommen, dass vier **OBSERVATIONEN** im Zeitintervall  $t=[1,4]$  generiert wurden. Diese vier **OBSERVATIONEN** tragen **SEMANTIK**, **INHALT** und **IDENTIFIKATION** und werden als ersten Inferenzschritt in der Observationsmenge gespeichert, welche als Basis für die Aktualisierung des dynamischen **SB-MODELLS** dient.

Der zweite Schritt ist die initiale Erstellung und Aktualisierung des Modells. Die **OBSERVATION 1** und **3** ist mit dem **SB-KNOTEN** „Flugzeughöhe“ assoziiert und aktivieren diesen Knoten für die **IDENTIFIKATION** „AC1“ und „AC2“. Die zweite Fixation generiert zwei **OBSERVATIONEN** 2.1 und 2.2, welche jeweils mit dem Knoten „Flugzeugposition“ und „Flugzeugkurs“ assoziiert sind. Die **OBSERVATION 4** aktiviert die „Position Flugverbotszone“.

Als dritter Schritt werden nun aggregierte Knoten aktiviert. Hier gibt es die „Höhentrennung“, die auf den zwei **SB-KNOTEN** „Flugzeughöhe“ aufbaut. Aus den **SB-KNOTEN** „Flugzeugkurs“ und „Flugzeugposition“ wird die „Projektion der Flugzeugposition“ generiert, da beide Elternknoten die gleiche **IDENTIFIKATION** AC2 haben. Außerdem wird der **SB-KNOTEN** „Distanz Flugverbotszone“ aktiviert, da es zwei Elternknoten „Flugzeugposition“ und „Position Flugverbotszone“ mit unterschiedlichen **IDENTIFIKATIONEN** gibt. Es gilt noch zu bemerken, dass die jeweiligen Observations durch das Netz weitergegeben werden. So ist beispielsweise die Observation 1 und 3 dem **SB-KNOTEN** „Höhentrennung“ zugeordnet. Ein anderes Beispiel ist der Knoten „Distanz Flugverbotszone“, der mit den **OBSERVATIONEN** der Elternknoten „Flugzeugposition“ und „Position Flugverbotszone“ verknüpft ist. Diese Observations werden im weiteren Verlauf für die Abweichungsberechnung zwischen Knoten verwendet.

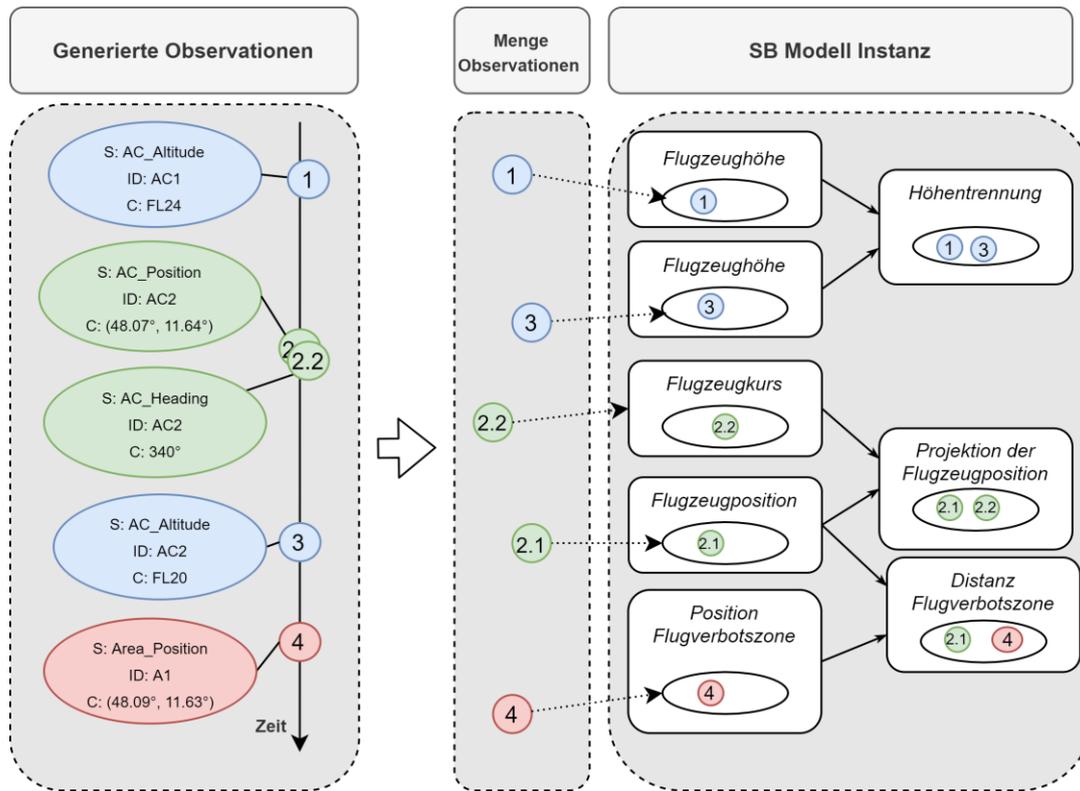


Abbildung 3-9 - Beispiel einer Inferenz basierend auf dem Modell in Abbildung 3-8.

### Inferenz des Systemzustands „Ground Truth“

Bei der Inferenz des Systemzustands wird der gleiche Inferenzmechanismus wie für den Bediener verwendet. Der einzige Unterschied ist, dass die Eingangsdaten der Bewertung nicht aus der Aufmerksamkeitsanalyse stammen, sondern von einem „perfekten Beobachter“. Dieser Beobachter nimmt alle Interessensgebiete gleichzeitig wahr, insbesondere auch die Gebiete, die auf der Bedienoberfläche gerade nicht zu sehen sind. Das heißt beispielsweise, wenn ein Bediener nur einen bestimmten Ausschnitt der taktischen Karte anzeigt, nimmt der künstliche Beobachter trotzdem alle Informationen auf der Karte wahr. Die Bewertung bekommt somit kontinuierlich jede mögliche Observation aus dem System mit dem aktuellen Inhalt. Das daraus abgeleitete Stand des Situationsbewusstseinsmodell wird als das „richtige“ Situationsbild abgelegt.

### 3.3.2 ABWEICHUNGSBERECHNUNG

Die Instanzen von Bediener und System bauen auf dem gleichen Modell auf und werden verglichen, um fehlerhaftes und fehlendes SB zu erkennen. Beide Modelle können aktive **SB-KNOTEN** halten und jeder dieser aktiven Knoten enthält eine Menge von mindestens einer zugehörigen **OBSERVATION**, wobei der **INHALT** der **OBSERVATIONEN** möglicherweise heterogen in Bezug auf den Datentyp ist. Ein **SB-KNOTEN** mit zwei **OBSERVATIONEN** könnte zum Beispiel jeweils einen **INHALT** mit Positions- und einen mit Zeitdaten halten. Die Abweichung zweier **SB-KNOTEN** kann nur durch den Vergleich auf der Ebene der **OBSERVATIONEN** zwischen Bediener- und Systemmodell quantifiziert werden. **OBSERVATIONEN** mit gleicher **SEMANTIK** tragen gleiche Datenvarianten als **INHALT**. Für eine generische Berechnung von Unterschieden zwischen zwei Observationen wird eine Abweichungsfunktion in Form der folgenden Exponentialfunktion definiert:

$$x_{dev}(t) = 1 - e^{-\frac{|dev(c_B(t),c_S(t))|}{c_n}} \quad (1)$$

Hier sind  $c_B$  und  $c_S$  die Werte des **INHALTS** der Bediener- beziehungsweise Systemknoten. Die Funktion  $dev(c_1, c_2)$  beschreibt die Abweichung beider Varianten und muss datentypspezifisch definiert werden, wie z.B. die Großkreisdistanz für zwei Koordinaten oder die Dauer zwischen zwei Zeitpunkten (siehe 3.5.2.4 für eine konkrete Implementierung). Die Konstante  $c_n$  ist ein Normalisierungsfaktor, welcher im statischen **SB-MODELL** für jede **OBSERVATION** festgelegt wird. Dieser frei gewählte Parameter skaliert das Ergebnis der Abweichungsfunktion und bestimmt, welcher Abstand als wesentlich modelliert wird. Die Abweichungsfunktion nimmt stets einen Wert zwischen 0 und 1 ein.

Wie in Abbildung 3-9 dargestellt, können **SB-KNOTEN** mehrere **OBSERVATIONEN** halten, deren Abweichung je nach Relationstyp zu den Elternknoten berechnet wird. Für AND- sowie DIFF-Relationen werden die einzelnen Abweichungswerte multipliziert:

$$x_{dev}^{node}(t) = \prod_i^n x_{dev}^i(t) \quad (2)$$

Für OR-Relationen wird der gesamte Abweichungswert durch den kleinsten Abweichungswert bestimmt:

$$x_{dev}^{node}(t) = \min(x_{dev}^i(t)) \quad (3)$$

Das vorangegangene Beispiel wird in Abbildung 3-10 nochmals aufgegriffen und die Abweichungsberechnung für den **SB-KNOTEN** „Höhentrennung“ beschrieben.

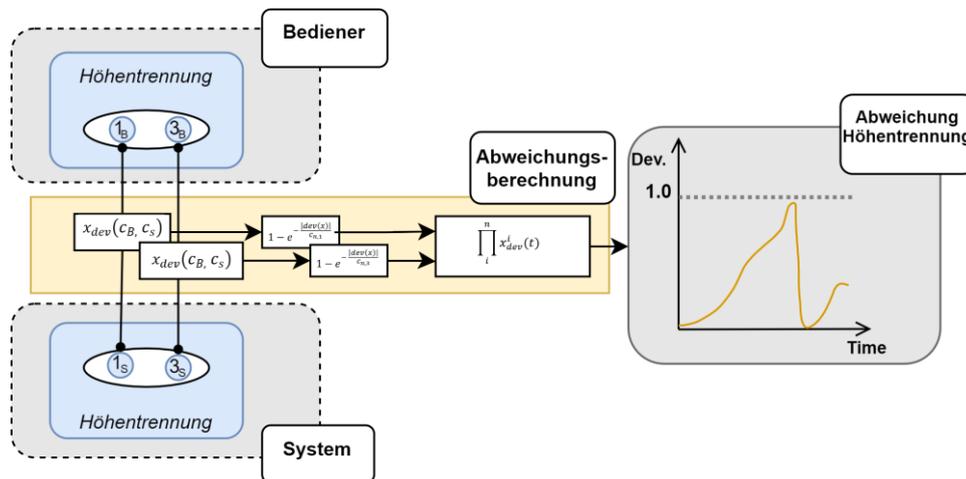


Abbildung 3-10 - Berechnung einer Abweichung im Knoten „Höhentrennung“ zwischen den Observation 1 und 3 mit exemplarischem Zeitverlauf der Abweichungsberechnung.

Der Knoten hält eine Menge von zwei **OBSERVATIONEN**, in diesem Fall „Flughöhe“ für das AC1 und „Flughöhe“ für das AC2. Beide **OBSERVATIONEN** im Bedienermodell werden mit der entsprechenden **OBSERVATION** im Systemmodell verglichen, indem die Abweichungsfunktion angewendet wird. Bei der Flughöhe liefert eine einfache Subtraktion der skalaren Werte das Ergebnis, während bei den Positionsinformationen ein Maß für die räumliche Trennung der beiden Inhalte genutzt werden kann.

Für die Umsetzung muss also eine sinnvolle Abweichungsfunktion für die verwendeten **INHALTE** definiert werden (siehe Abschnitt 3.5.2.4).

#### Sonderfall „Unwissenheit“

Die beschriebene Funktion zur Abweichungsberechnung basiert bisher darauf, dass beim Bediener alle notwendigen Observationen mindestens einmal gemessen wurden. Dies ist der Fall bei vielen Systeminformationen von deren Existenz ein Bediener gesichert weiß, wie beispielweise ein Pilot sich im Klaren darüber ist, dass sein Flugzeug auf einer bestimmten Flughöhe fliegt. Der bereits angesprochene Informationsraum vieler Arbeitsplätze umfasst aber auch Informationen über Objekte, von deren Existenz der Bediener nicht unbedingt weiß wie beispielsweise ein unbekanntes Flugzeug im Luftraum. Um die Fälle abzubilden, bei denen das System bereits von diesen Informationen weiß, der Bediener sich deren Existenz aber nicht bewusst ist, wird ein Sonderfall der Abweichungsberechnung eingeführt. Sollten die notwendigen Informationen im System- aber nicht im Bedienermodell vorhanden sein, wird die Abweichungsfunktion auf  $x_{dev}(t) = -1.0$  gesetzt.

#### Limitierungen der Modellierung

Die hier beschriebene Modellierung trifft Vereinfachungen, die die Bewertung einschränken und sollen hier nochmal explizit aufgeführt werden:

- Das Modell beinhaltet keinen Mechanismus, der das „Vergessen“ einer Information modelliert.
- Das Bewertungsverfahren nimmt an, dass jede fixierte Information auch wahrgenommen wurde und kann damit keine Aussage über die Unsicherheit der zugrundeliegenden Aufmerksamkeitsanalyse geben.
- Der logische Graph hat keinerlei Möglichkeiten zur Integration von dynamischem Systemverhalten. Mentale Modelle, die der internen Repräsentation des Bedieners vom System entsprechen, können hier nur statisch definiert werden im Sinne der involvierten Informationen aber nicht Sinne eines zeitdynamischen Verhaltens.

Wie im Folgenden beschrieben wird, kann das vorgestellte Messsystem trotz dieser Limitierungen nützliche Aussagen über das SB seines Bedieners treffen. Die Vereinfachungen wurden zugunsten einer einfachen Modellierung und Validierung getroffen und stellen eine gute Ausgangslage zur Umsetzung des Systems dar.

### 3.4 INTERVENTIONSGENERIERUNG

Im Folgenden wird beschrieben, wie die maschinelle Messung des SB als Basis für ein Assistenzsystem genutzt werden kann, um gezielt Interventionen zu steuern und das in Abschnitt 3.1 geforderte Verhalten umzusetzen. Dort heißt es, dass das Assistenzsystem den Bediener auf nicht wahrgenommene, aber relevante Informationen hinweisen soll. Der Adaptionsmechanismus benötigt

hierfür die SB-Bewertung, eine Analyse der aktuellen Situation sowie Wissen über die aufgabenrelevanten Informationen. Dieser Aufbau ist in Abbildung 3-11 dargestellt.

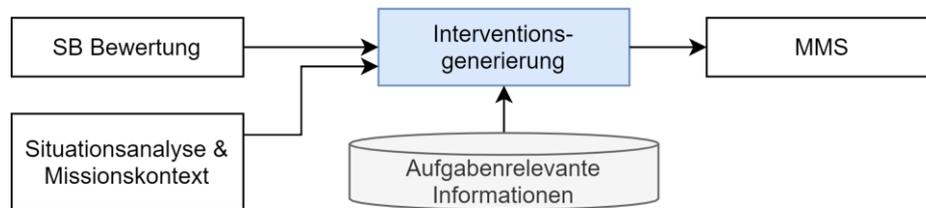


Abbildung 3-11 - Adaptionsmechanismus mit den Eingangsgrößen SB-Bewertung sowie der Analyse von Missionskontext und Situation. Die wirkt auf die MMS.

Das Verhalten der Assistenz kann nun in Bezug auf das beschriebene **SB-MODELL** wie folgt konkretisiert werden: Die Interventionsgenerierung analysiert kontinuierlich, welche Informationen in der aktuellen Situation relevant sind. Dazu braucht es einerseits Informationen über die aktuelle Situation und andererseits Wissen, welche Informationen für die aktuelle Aufgabensituation relevant sind. Vorausgesetzt diese Informationen sind als **SB-KNOTEN** modelliert, wird der Abweichungswert aller relevanten **SB-KNOTEN** ermittelt. Sofern die Abweichung unter einem definierten Grenzwert ist, wird keine Intervention ausgelöst. Ist die Abweichung in einem der Knoten zu hoch, wird für einen festgelegten Zeitraum gewartet, bevor eine Intervention ausgelöst wird. Nachdem eine Intervention ausgelöst wurde, wird der **SB-KNOTEN** des Systems zu diesem Zeitpunkt gespeichert. Um zu verhindern, dass zu viele Interventionen über die gleiche Information kommen, wird geprüft, ob sich die **INHALTE** dieses Knotens seit der letzten Intervention geändert hat. Falls dem so ist, sind neue Interventionen möglich.

### 3.4.1 GESTALTUNG DER INTERVENTION

Grundsätzlich sind die in der vorliegenden Arbeit verwendeten Interventionstypen auf zwei Kategorien beschränkt. **Lenkung von Aufmerksamkeit** auf bereits dargestellte Informationen in der Bedienschnittstelle und **Vermittlung von Zusatzinformation** über Sprachnachrichten oder Annotationen von Elementen der Bedienoberfläche. Eine Erweiterung der Eingriffsstrategien ist denkbar und wird im Ausblick dieser Arbeit nochmals aufgegriffen. Arbeiten wie beispielsweise die von Brand und Schulte (2018) bedienen ein breiteres Spektrum von Eingriffen in den Arbeitsprozess beispielsweise durch Vereinfachung oder Übernahme von Aufgaben durch die Automation.

### 3.4.2 IDENTIFIKATION VON RELEVANTEN INFORMATIONEN

Bei der Durchführung von Laborversuchen mit definierten Versuchsaufgaben ist die Wahl der Informationen meist trivial, da sie durch die Experimentalaufgabe vorgegeben wird. Im Falle von komplexerer Arbeitssituationen ist die Frage nach den situativ relevanten Informationen jedoch schwerer zu beantworten. Der Abschnitt 5.2 beschreibt ein Vorgehen, um Anwendungsfälle für die adaptive Unterstützung des Piloten zu identifizieren.

## 3.5 GESAMTSYSTEMARCHITEKTUR

Im finalen Systemaufbau werden die drei Module Aufmerksamkeitsanalyse, Situationsbewusstseinsbewertung und Interventionsgenerierung, wie in Abbildung 3-12 dargestellt, zu einer Gesamtsystemarchitektur zusammengefügt.

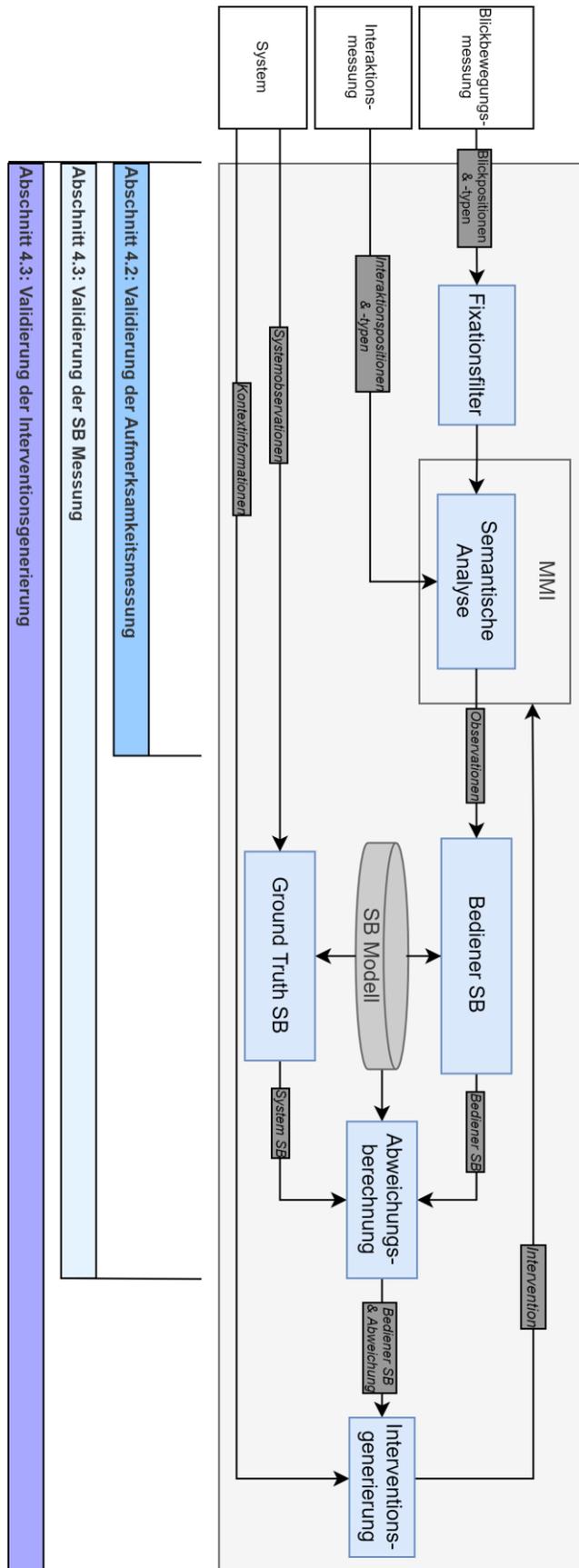


Abbildung 3-12 - Gesamtsystemarchitektur des SB-adaptiven Assistenzsystems.

### 3.5.1 AUFBAU

Die technische Umsetzung des Konzepts stellt folgende Anforderungen: Zunächst ist die Einrichtung von robuster Messtechnik für Blickbewegung und Interaktion notwendig, welche ein ausreichend hohes Maß an Genauigkeit bereitstellt, um zwischen kleinen Interessensgebieten auf der Bedienoberfläche zu unterscheiden. Zusätzlich ist für die semantische Analyse ein hohes Maß an Bedienschnittstellenintegration notwendig, um basierend auf Pixelkoordinaten auswerten zu können, welche Informationen dort dargestellt sind. Dies ist nicht möglich, wenn es keinen tieferen Eingriff in das Softwaresystem zur Oberflächendarstellung gibt, beziehungsweise keine Schnittstelle vorliegt, die dieses Wissen bereitstellt. Zuletzt muss eine Wissensbasis geschaffen werden, die das Domänenwissen der Arbeitsumgebung integriert, was in der vorgestellten Architektur an zwei Stellen besonders relevant ist. Zum einen muss das **SB-MODELL** relevante Informationen modellieren und Beziehungen zwischen Informationen korrekt darstellen. Andererseits muss die Situationsanalyse dynamisch feststellen können, welche Situation gerade vorliegt, um die relevanten Informationen abzuleiten.

### 3.5.2 IMPLEMENTIERUNG

Dieser Abschnitt beschreibt die softwaretechnische Implementierung des Assistenzsystems. Hier wird die Implementierung unabhängig von Peripheriegeräten sowie der Arbeitsplatzumgebung beschrieben. Die Integration dieser Implementierung in die Versuchsumgebung eines Cockpitsimulators wird in Kapitel 4.1 erläutert. Die folgende Implementierung wurde mit der Programmiersprache C++ umgesetzt unter Zuhilfenahme der Bibliotheken Boost, Eigen und Qt<sup>4</sup>.

#### 3.5.2.1 Observationen

Abbildung 3-12 zeigt, dass die zentrale Schnittstelle zwischen semantischer Analyse und Situationsbewusstseinsmodell die **OBSERVATIONEN** sind. Wie im Konzept in 3.2.1 dargelegt, tragen **OBSERVATIONEN** Informationen über **SEMANTIK**, **INHALT** und **IDENTIFIKATION**. In der entwickelten Software wurden die **OBSERVATIONEN** wie in Abbildung 3-13 modelliert.

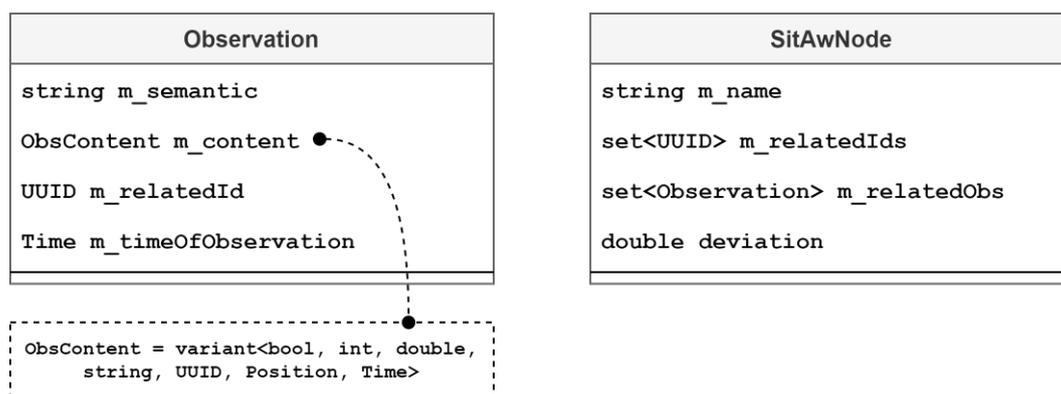


Abbildung 3-13 - Pseudoklassendiagramm einer `Observation` (links) und einer `SitAwNode` (rechts)

Hervorzuheben ist hier die Modellierung des Inhalts, welcher mittels einer `variant`-Datenstruktur modelliert wurde. Diese kann verschiedene zugrundeliegende Datentypen einnehmen, was in diesem Fall hilfreich ist, um unterschiedliche Arten von Inhalten zu modellieren. Die Variable `m_content` kann also boolesche Zustände, Ganzzahl, Gleitkommazahlen, Kennung, Positionen oder Zeitdaten

<sup>4</sup>Website: Boost (<https://www.boost.org/>), Eigen (<https://eigen.tuxfamily.org/>), Qt (<https://www.qt.io/>)

repräsentieren. Die Software für das MMS muss bei gegebener Fixationsposition in der Lage sein, diese Daten zu generieren.

### 3.5.2.2 SB-Modell

Die zugrundeliegende Datenstruktur der **SB-MODELLE** ist der `SitAwNode`, der einen **SB-KNOTEN** des **SB-MODELLS** repräsentiert. Ein Knoten trägt einen Namen sowie zwei Sets, welche alle assoziierten Identifikationen sowie alle assoziierten **OBSERVATIONEN** beinhalten. Darüber hinaus beinhaltet der Knoten noch einen Gleitkommawert, der die Abweichung dieses Knotens repräsentiert. Die Abbildung 3-13 zeigt das zugehörige Klassendiagramm. Das vordefinierte **SB-MODELL** `SitAwMap` besteht aus einem Feld von `SitAwNodes` und drei Transitionsmatrizen, die **RELATIONEN** zwischen den Knoten darstellen. Die drei Matrizen repräsentieren jeweils einen Kantentypen. Darüber hinaus beinhaltet die `SitAwMap` für jeden Knoten ein Feld von **OBSERVATIONEN** (`m_obsPerNode`), welches die Knoten mit **OBSERVATION** einer bestimmten Semantik assoziiert.

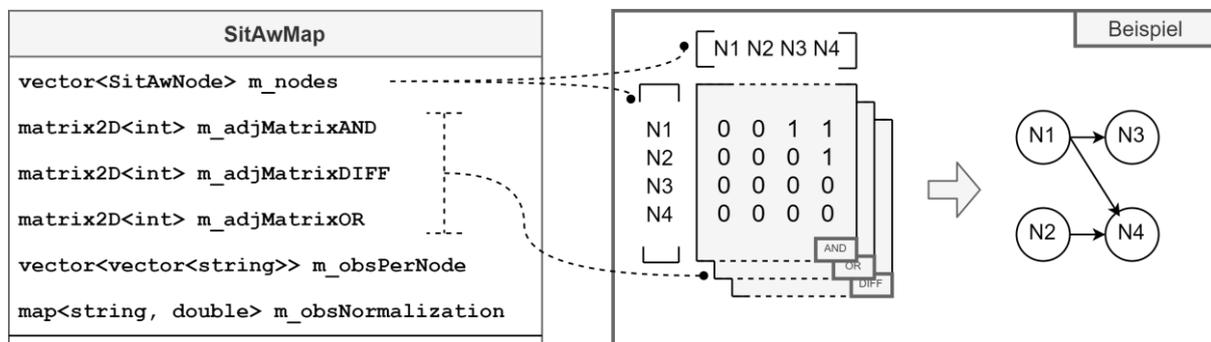


Abbildung 3-14 - Klassendiagramm der `SitAwMap` (links). Beispiel wie mithilfe der Transitionsmatrizen ein Netz modelliert werden kann (rechts).

### 3.5.2.3 SB Bediener- und Systemmodell

Zur Berechnung eines dynamischen **SB-MODELLS** auf Basis von **OBSERVATIONEN** dient die Klasse `SitAwSimulation`, deren Aufbau in Abbildung 3-15 dargestellt ist. Sie beinhaltet alle aktiven **SB-KNOTEN**, die statische Wissensbasis sowie alle aktuellen **OBSERVATIONEN**.

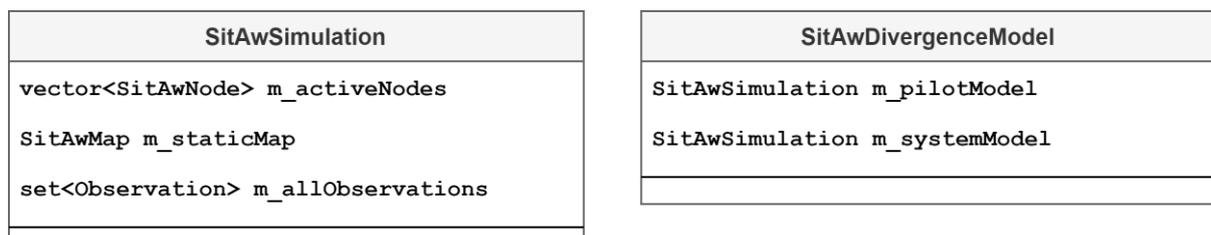


Abbildung 3-15 - Klassendiagramm der Komponenten zur dynamischen Berechnung der SB-Modelle

Der Algorithmus 3-1 zeigt den Pseudocode zur Aktualisierung der `SitAwSimulation`. Die Aktualisierung basiert auf dem aktuellen Observation-Set. Zunächst werden für alle Observationen aus `m_allObservations` die assoziierten **SB-KNOTEN** im SB-Modell `m_staticMap` gesucht. In allen gefundenen Knoten werden dann die Observationen im `m_relatedObs` aktualisiert. Nach der Aktualisierung wird für jeden Knoten überprüft, ob Kantenbedingungen erfüllt sind. Sofern eine Kantenbedingung erfüllt ist, wird der Kindknoten entweder erstellt oder aktualisiert. Dieser Schritt wird so lange ausgeführt, bis es keine Veränderungen mehr im Graphen gibt und damit die Konvergenzbedingung erfüllt ist.

SITAW-UPDATE (OBSERVATION-SET, SBMODEL)
<pre> FOR <b>OBS</b> IN <b>OBSERVATION-SET</b>:     <b>NODES</b> ← FINDE MIT <b>OBS</b> ASSOZIIERTE <b>SITAWNODES</b> IM <b>SBMODEL</b>     FOR <b>NODE</b> IN <b>NODES</b>:         IF <b>NODE</b> EXISTIERT:             AKTUALISIERE <b>RELATEDOBS</b> IN <b>NODE</b> MIT <b>OBS</b>         ELSE:             ERSTELLE NEUE <b>SITAWNODE</b>  <b>WHILE</b> NICHT KONVERGIERT:     <b>NODES</b> ← ALLE STATISCHEN <b>SITAWNODES</b> IN <b>SBMODEL</b>     FOR <b>NODE</b> IN <b>NODES</b>:         <b>PARENTS</b> ← FINDE AKTIVE PARENTNODES         IF KANTENBEDINGUNG DURCH <b>PARENTS</b> ERFÜLLT:             IF <b>CHILD-SITAWNODE</b> EXISTIERT:                 AKTUALISIERE <b>CHILD-SITAWNODE</b> MIT <b>RELATEDOBS</b> DER <b>PARENTS</b>             ELSE:                 ERSTELLE NEUE <b>CHILD-SITAWNODE</b> </pre>

Algorithmus 3-1 Algorithmus zur Aktualisierung der dynamischen SB-Modellinstanzen für Bediener und System.

Die Modellinstanzen des Bediener und des Systems werden beide im `SitAwDivergenceModel` berechnet. Diese Anwendung empfängt kontinuierlich die gemessenen Pilot- sowie Systemobservationen. Mit diesen wird ein Set von **OBSERVATIONEN** (`m_allObservations`) befüllt, welche stets die aktuelle Version einer **OBSERVATION** beinhaltet. Gleichheit von **OBSERVATIONEN** ist definiert als Gleichheit der **SEMANTIK** und der **IDENTIFIKATION**. Das `SitAwDivergenceModel` stellt Funktionen bereit, mit deren Hilfe sich die Abweichung in spezifischen Knoten abfragen lässt.

### 3.5.2.4 Abweichungsberechnung

Das Konzept zur Abweichungsberechnung setzt eine Abweichungsfunktion  $dev(x_1, x_2)$  voraus, die die absolute Distanz zweier Inhalte des gleichen Datentyps berechnet. Die Tabelle 3-1 zeigt Abweichungsfunktionen für die verschiedenen Datentypen Ganz- und Gleitkommazahl, Text, Wahrheitswert, Zustände, Position und Zeit.

Informationstyp	Abweichungsfunktion $dev(x_1, x_2)$	Beispiel
Numerische Werte Ganz- & Gleitkommazahl	$x_2 - x_1$	Höhe, Geschwindigkeit, Temperatur
Text	$\begin{cases} 0.0 & \text{if } x_1 \equiv x_2 \\ 1.0 & \text{otherwise} \end{cases}$	Textnachricht
Wahrheitswert		Landelampe an/aus
Zustände / Modi		Autopilot ALT Hold/Acquire
Position	Großkreisdistanz [m] von $x_1$ zu $x_2$	Gebäudeposition
Zeit	Dauer [s] zwischen Zeitpunkt $x_1$ und $x_2$	Uhrzeit

Tabelle 3-1 - Abweichungsfunktionen der verschiedenen Grunddatentypen

Softwaretechnisch ist die Auswahl der richtigen Abweichungsfunktion für zwei Objekte des Typs `ObsContent` (siehe Abbildung 3-13) mithilfe des Visitor-Patterns der C++-Bibliothek *Boost* implementiert.

### 3.5.2.5 Interventionsgenerierung

Die Interventionsgenerierung erhält den aktuellen Stand des `SitAwDivergenceModel` als Wissensbasis für die Interventionsgenerierung. Ferner erhält sie Kontextinformationen `Context`, auf deren Basis eingeschätzt wird, welche Informationen gerade relevant sind. Der `Context` muss auf Abfrage alle relevanten `SitAwNodes` der vorliegenden Aufgabensituation bereitstellen. Mithilfe dieser Eingangsvariablen generiert der Algorithmus 3-2 informationsspezifische Interventionen.

ADAPTIVE-INTERVENTION (CONTEXT, SITAWDIVMODEL)
<b>WHILE ASSISTENZ-AKTIV:</b> <b>NODES</b> ← ALLE RELEVANTEN INFORMATIONEN IM AKTUELLEN <b>CONTEXT</b> <b>FOR RELEVANT-INFO IN NODES</b> <b>TIME-UNTIL-NOTIFICATION</b> ← SPEZIFISCHE ZEITKONSTANTE FÜR <b>RELEVANT-INFO</b> TRIGGER-INTERVENTION( <b>RELEVANT-INFO</b> , <b>SITAWDIVMODEL</b> , <b>TIME-UNTIL-NOTIFICATION</b> )
TRIGGER-INTERVENTION (RELEVANT-INFO, SITAWDIVERGENCEMODEL, TIME-UNTIL-NOTIFICATION)
<b>TIME-OF-HIGH-DEV</b> ← LETZTER ZEITPUNKT HOHER ABWEICHUNG VON <b>RELEVANT-INFO</b> <b>DEVIATION</b> = <b>SITAWDIVERGENCEMODEL</b> .COMPUTE-DEVIATION ( <b>RELEVANT-INFO</b> ) <b>IF DEVIATION</b> > 0.6 <b>NOW</b> ← AKTUELLE ZEIT <b>IF TIME-OF-HIGH-DEVIATION</b> == NIL <b>TIME-OF-HIGH-AD</b> ← <b>NOW</b> <b>IF (NOW – TIME-OF-HIGH-DEVIATION)</b> > <b>TIME-UNTIL-NOTIFICATION</b> TRIGGER-INTERVENTION ( <b>INFORMATION</b> ) <b>TIME-OF-HIGH-AD</b> ← NIL <b>ELSE</b> <b>TIME-OF-HIGH-AD</b> ← NIL

Algorithmus 3-2 – Interventionsgenerierung basierend auf Kontext und SB-Abweichungsmodell.

Für alle relevanten Informationen wird überprüft, ob diese seit einer vordefinierten Zeit über einem Abweichungslimit von  $x_{dev}(x_1, x_2) > 0.6$  liegen. Der Abweichungswert ist durch Parameter  $c_n$  skaliert, wobei bei einer Abweichung von  $dev(x_1, x_2) = c_n$  ein Abweichungswert von  $x_{dev}(x_1, x_2) = 0.63$  resultiert. Der gewählte Grenzwert von 0.6 entspricht damit ungefähr einer absoluten Abweichung von  $c_n$ . Eine Wartezeit bis zur Intervention ist grundsätzlich notwendig, um dem Bediener auch Zeit zu lassen, die Information ohne Lenkung seiner Aufmerksamkeit wahrzunehmen. Sofern die Abweichung länger als die Zeitkonstante über dem Grenzwert liegt, wird eine Intervention ausgelöst. Die spezifische Gestaltung der Intervention ist arbeitsplatzspezifisch und wird für die Anwendung in dieser Arbeit im Abschnitt 4.4.1.1 beschrieben.

## 4 VALIDIERUNGSSTUDIEN

Das vorliegende Kapitel beschreibt drei Studien, die zur Validierung des SB-adaptiven Assistenzsystems durchgeführt worden sind. Die Versuchsumgebung der Studien war ein Kampfflugzeugsimulator des Instituts für Flugsysteme (IFS) der Universität der Bundeswehr München (UniBwM) durchgeführt worden sind.

Die durchgeführten Studien werden in der Reihenfolge Aufmerksamkeitsanalyse, Situationsbewusstseinsbewertung und Interventionsgenerierung vorgestellt. An dieser Stelle sei auf Abbildung 3-12 verwiesen, die zeigt, welchen Teil des Gesamtsystems die drei Validierungsversuche untersuchen. Die folgende Tabelle zeigt einen Versuchsüberblick mit Teilnehmergruppe, Probandenanzahl sowie die gemessenen Variablen.

Fokus der Studie	Teilnehmergruppe	Probanden- zahl	Versuchs- anzahl	Gemessene Variablen
Aufmerksamkeits- abschätzung (4.2)	Piloten mit Fluglizenz	10	2	Blickposition Abstand zu Interessensgebiete Dauer Fixation Dauer Fixation-Reaktion Dauer Aufgabenbearbeitung
Situationsbewusstseins- bewertung (4.3)	Kampfflugzeugpiloten der dt. Luftwaffe  Anwärter fliegerischer Dienst der dt. Luftwaffe	8  9	2	SB-Abweichung SART Spez. Leistungsmaße: <u>Versuch 1</u> : Tiefflugleistung <u>Versuch 2</u> : Bearbeitungszeit
Interventionsgenerierung (4.4)	Piloten mit Fluglizenz	10	2	SB-Abweichung Interventionsanzahl Erkennungsdauer Fragebogen (Bewertung Assistenz) Spez. Leistungsmaße: <u>Versuch 1</u> : Fehlerrate Verfolgung Flughöhe, -geschwindigkeit und -pfad <u>Versuch 2</u> : Delegationszeit

Tabelle 4-1 - Überblick über die durchgeführten Validierungsstudien mit Teilnehmergruppen, Probanden- und Versuchszahl sowie den abhängigen Variablen

Bevor die Einzelstudien beschrieben werden, stellt der erste Abschnitt 4.1 den in allen Studien verwendeten Versuchsaufbau und Cockpitsimulator vor. Darauf folgend wird im Abschnitt 4.2 der Ansatz der fixationsbasierten Analyse von Aufmerksamkeit auf Interessensgebiete im Cockpit validiert. Dieser Abschnitt steht im Zusammenhang mit der Publikation (Schwerd & Schulte, 2022). Der Abschnitt 4.3 stellt die Studie zur Validierung der SB-Bewertung vor (Schwerd & Schulte, 2020) und Abschnitt 4.4 beschreibt einen Versuch zur Erprobung der Interventionsgenerierung auf Basis der SB-Bewertung (Schwerd & Schulte, 2021c).

### 4.1 VERSUCHSAUFBAU

Im folgenden Abschnitt wird die die Versuchsumgebung vorgestellt, welche in allen Studien zum Einsatz kam. Zunächst wird die Gestaltung des Cockpits vorgestellt mit allen Bedien- und Anzeigeschnittstellen. Dann werden die verwendeten Messsysteme beschrieben und zuletzt die technische Umsetzung des Simulationssystems.

#### 4.1.1 SIMULATORCOCKPIT

Der Cockpitsimulator des IFS ist in Abbildung 4-1 dargestellt. Das Cockpit besteht aus einem zentralen und zwei seitlichen Touchscreens sowie einer gekrümmten Außensichtdarstellung der Firma Q4 mit einem horizontalen Sichtfeld von 210° und einem vertikalen Sichtfeld von 45°<sup>5</sup>. Zur Kontrolle des Flugzeugs dient das HOTAS-System von Reiser Systemtechnik bestehend aus Schubhebel und Steuerknüppel<sup>6</sup>, welche beide das im Eurofighter Typhoon verbaute System nachbildet. Zusätzlich hierzu sind noch Ruderpedale verbaut. Die Bedienoberflächen stellen Systeminformationen, eine taktische Karte sowie Schnittstellen für das Missionsmanagement und die Führung der *Unmanned Aerial Vehicles* (UAV) dar. Alle Bedien- und Anzeigeelemente des Cockpits werden im Abschnitt E des Anhangs vorgestellt.



Abbildung 4-1 - Der Cockpitsimulator am IFS besteht aus einem Sichtsystem mit HUD, drei Touchscreens sowie einem HOTAS System und Ruderpedalen zur Bedienung.

#### 4.1.2 MESSTECHNIK

Die zentralen Messverfahren waren die Blickbewegungsmessung und die Interaktionsmessung, welche im folgenden Abschnitt beschrieben werden. Zur Blickbewegungsmessung wurde ein kommerzielles Messsystem integriert und zur Integration wurde die vorhandene Software zur Bedienung der HOTAS und Touchscreens verwendet.

##### Blickbewegungsmessung

Zur Messung der Blickbewegung wurde das kamerabasierte Messsystem der Firma SmartEye<sup>7</sup> in das Cockpit integriert. Das Messsystem nutzt die videobasierte Auswertung der Korneareflection in

<sup>5</sup> <https://q4services.com/q4-directvue-direct-dome-projection/>

<sup>6</sup> <https://www.reiser-st.com/components-2/eurofighter-typhoon-hotas/>

<sup>7</sup> SmartEyePro 0.3MP (<https://smarteys.se/>)

Kombination mit der Abschätzung der Kopfposition und -rotation (Duchowski, 2017). Hierfür wurden im Cockpit vier Kameras und drei Infrarotquellen platziert. Das Messsystem legt sich mittels Bilderkennung von Gesichtsmerkmalen ein Kopfmodell an und kann nach einem Kalibrierungsverfahren die Blickrichtung mit 60Hz abschätzen.

In der von SmartEye bereitgestellten Software ist ein dreidimensionales „Weltmodell“ zu definieren, das den Cockpitaufbau inklusive der Messinstrumente beinhaltet. Abbildung 4-2 zeigt schematisch das verwendete Modell mit der Position der Kameras und Infrarotquellen. Das Modell besteht aus einem *Multi-Functional-Head-Down-Display* (MHDD), zwei *Side-Head-Down-Displays* (SHDD) sowie dem *Head-Up-Display* (HUD). Wenn der Blick ein Display im Weltmodell schneidet, werden die folgenden Daten über eine TCP/IP-Verbindung versendet: Zeitstempel, Pixelposition und Displayname des Schnittpunkts. Das Messsystem stellt auch eine Blickbewegungsklassifikation bereit, welche zwischen Fixation, Sakkade und Blinzeln unterscheidet. Diese Informationen werden mit einem Zeitversatz von  $t \approx 800ms$  über eine zweite TCP/IP-Verbindung geschickt.

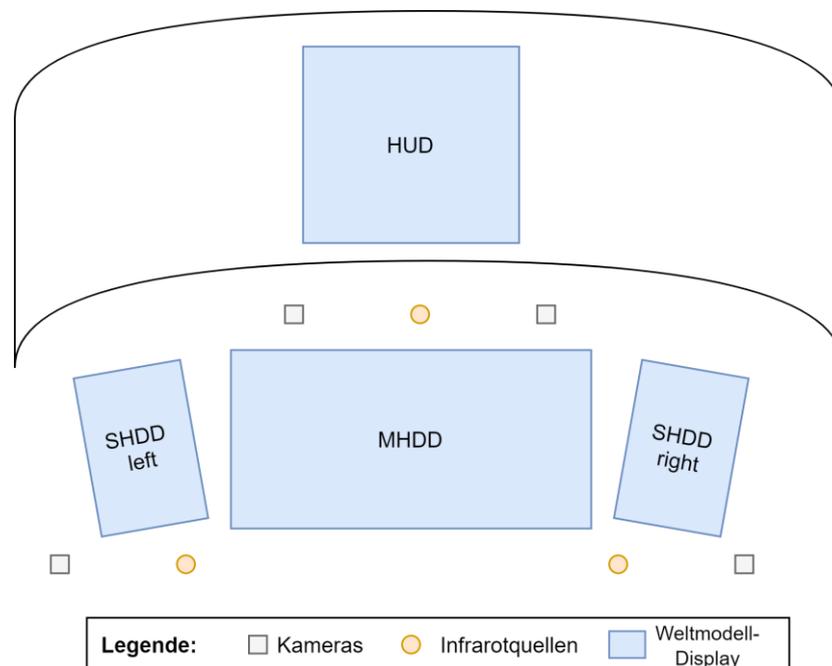


Abbildung 4-2 - Weltmodell des Blickbewegungsmesssystems inklusive Positionierung der Infrarotstrahler und Kameras.

## Interaktionsmessung

Das HOTAS-System ist über ein CAN-Bus an das Simulationssystem angeschlossen, welches mithilfe eines CAN2USB-Adapters ausgelesen wird. Die Bewegung des Steuerknüppels oder der Schubhebel werden in Nachrichten übersetzt und können sowohl zur Auslösung von Steueraktion als auch zur Messung von Interaktionen verwendet werden. Die Interaktionsmessung auf den Bedienoberflächen wurde softwareseitig umgesetzt und nutzt die Verarbeitung der Berührungspunkte mithilfe der Software Qt. Auch hier werden die manuellen Interaktionen gleichzeitig zur Bedienung und Messung weiterverarbeitet.

### 4.1.3 SIMULATIONSUMGEBUNG

Die Simulation besteht aus der Integration verschiedener Applikationen, die die virtuelle Welt simulieren, darstellen oder bedienbar machen. Die Abbildung 4-3 zeigt den Systemaufbau des Simulators mit einem Fokus auf die Umsetzung des SB-adaptiven Systems. Hier sind die Applikationen

in drei Kategorien unterteilt: Simulation/Automation (S/A), Mensch-Maschine-Schnittstelle (MMS) und Adaptive Assistenz (AA). Die Applikationen kommunizieren über ein IP-basiertes Netzwerk, was softwaretechnisch durch den Einsatz der Kommunikationsplattform *AnyCom* unterstützt wird.

Die Komponenten zur Simulation und Automation (S/A) bestehen zum einen Teil aus von Mitarbeitern des IFS entwickelten Komponenten wie beispielsweise *Scenario*, *Blue- und RedForce* sowie die Cockpit- und Waffensimulation. Zum anderen Teil wurden kommerzielle Werkzeuge eingesetzt, wie beispielsweise bei der Simulation des Kampfflugzeugs mithilfe von *X-Plane*.<sup>8</sup>

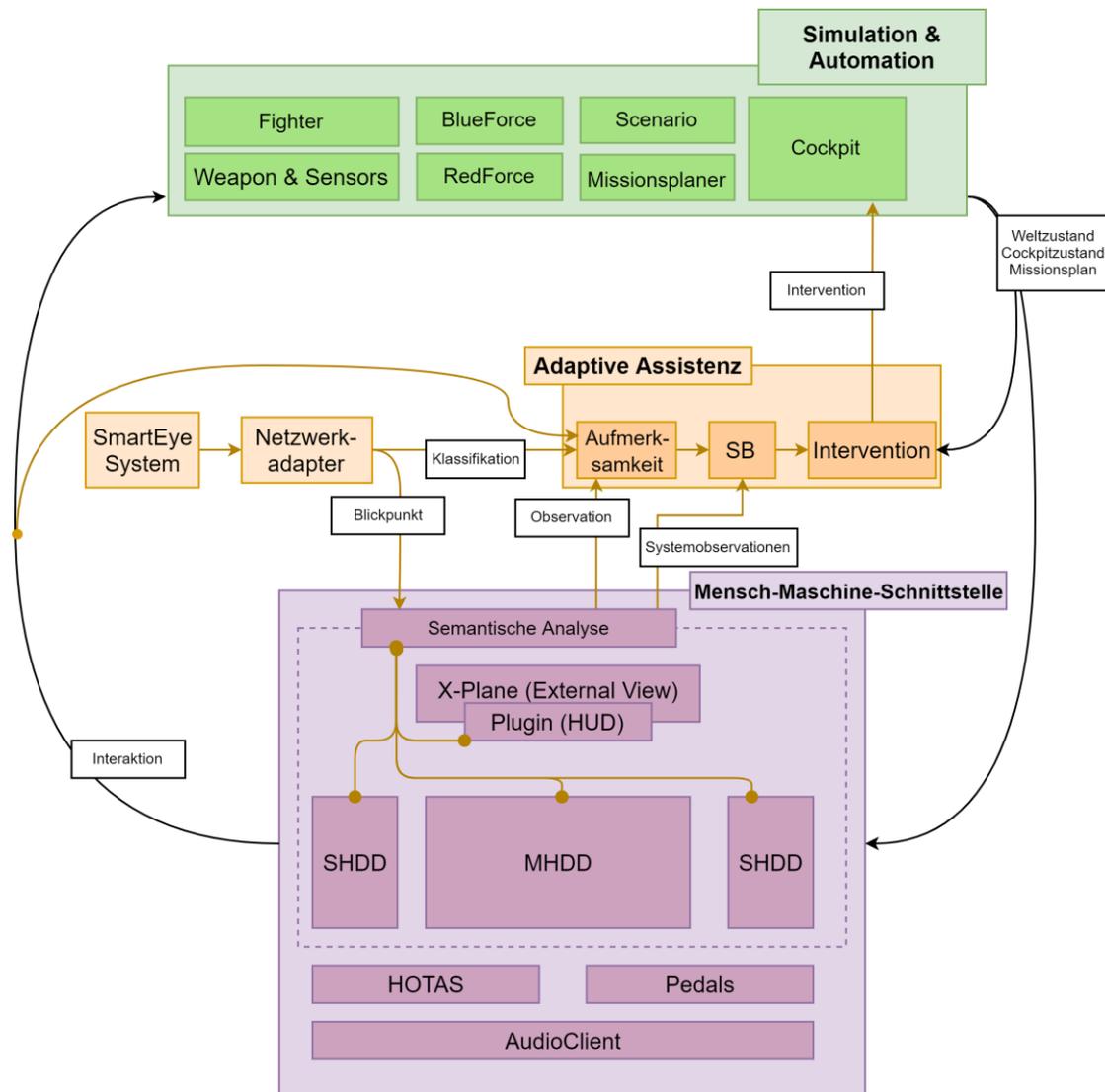


Abbildung 4-3 - Systemarchitektur der Simulationsumgebung bestehend aus Simulations- und Automationsapplikationen (grün), der MMS (lila) sowie der Applikationen für die adaptive Assistenz (orange).

Beim MMS ist die Bedienoberfläche (*SHDD* und *MHDD*) selbst entwickelt mithilfe der Softwarebibliothek *Qt* sowie der Kartendarstellungssoftware *Carmenta*<sup>9</sup>. Die Außensicht wurde mithilfe von *X-Plane* umgesetzt, ergänzt durch ein Plugin für das *HUD*. Die Applikationen *HOTAS* und *Pedals* basieren auf hardware-spezifischen Treibern zum Auslesen des CAN-Bus der Steuergeräte. Der *AudioClient* setzt die Funkkommunikation, die Warntöne sowie die *Text-to-Speech*-

<sup>8</sup> X-Plane 11 (<https://www.x-plane.com/>)

<sup>9</sup> Carmenta, <https://carmenta.com/>

Benachrichtigungen um. Die Funkkommunikation wurde mithilfe der Software *teamspeak*<sup>10</sup> realisiert und die Benachrichtigungen mittels einer Sprachsynthesoftware *Flite* der Carnegie Mello University implementiert (*CMU Flite: Speech Synthesizer*, 2018).

Die AA baut auf dem bereits beschriebenen Blickbewegungssystem von SmartEye auf. Die Implementierung wird in Abschnitt 3.5.2 beschrieben und wurde so integriert, dass die gemessenen Blickbewegungspunkte an das MMS geschickt werden und die *Semantische Analyse* die daraus erstellten **OBSERVATIONEN** an die AA weiterschickt. Gleichzeitig werden kontinuierlich Systemobservationen generiert, welche ebenfalls an das AA verschickt werden. Parallel hierzu werden die Klassifikationen der Blickbewegungspunkte direkt an die AA geschickt und dort mit den Observationen zusammengeführt, um diese nur nach fixierten Informationen zu filtern. Eine Besonderheit dieser Architektur ist, dass der Zustand der MMS in einer getrennten Applikation verwaltet wird. Aus diesem Grund gehen Interventionen der adaptiven Assistenz zuerst in das S/A-Modul, bevor sie von diesem an die Schnittstelle verteilt werden. Für diese Gestaltung gibt es zwei Gründe: Im Zusammenhang mit den Versuchen schafft es die Möglichkeit, aufgezeichnete Daten nochmal an die MMS zu geben und auf dieser Basis die semantische Auswertung zu verbessern. Der zweite Grund ist, dass es die Spiegelung des MMS auf einem zweiten Rechner ermöglicht. Dies stellte sich im Verlauf der Versuche als notwendig heraus, da sich die semantische Auswertung auf der taktischen Karte mit der Kartendarstellungssoftware, als rechenaufwendig erwies, vor allem im Zusammenhang mit der Generierung des Systemzustands. Die semantische Auswertung wurde dann auf der gespiegelten Version der Cockpitapplikation MHDD und SHDD durchgeführt. Die Applikation zur Anzeige des HUD war dagegen performant genug, um die semantische Auswertung live durchzuführen. Im Folgenden werden noch anwendungsspezifische Details der Implementierung beschrieben

### Auswertungsradius

Wie in Abschnitt 3.2.1 dargestellt, wird nicht nur der gemessene Fixationspunkt ausgewertet, sondern auch die nahe Peripherie des Blickpunkts. Dies wurde mithilfe eines probenbasierten Ansatzes umgesetzt. Hierfür wird um den Fixationspunkt herum in beide Raumdimension des Bildschirms Proben im Abstand von  $dx$  beziehungsweise  $dy$  generiert (siehe Abbildung 4-4). In der Implementierung im Rahmen der durchgeführten Studien galt  $dx = dy = 10px$ . Alle generierten Proben wurden genutzt, um **OBSERVATIONEN** zu generieren, was im folgenden Abschnitt beschrieben wird.

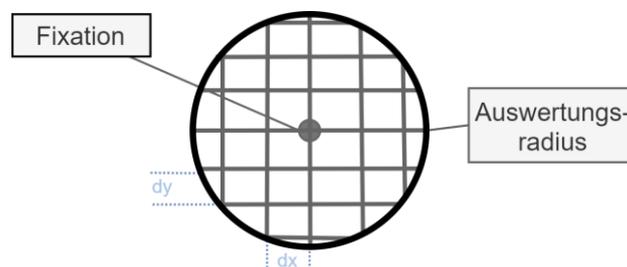


Abbildung 4-4 - Probenbasierter Ansatz zur Auswertung des Radius um eine gemessene Fixation. Proben werden an den Schnittpunkten der grauen Linien durchgeführt.

<sup>10</sup>Teamspeak, <https://www.teamspeak.com/en/>

## Generierung von Piloten- & Systemobservationen

Eine zentrale Funktion der SB-Bewertung ist die semantische Analyse, also die Generierung von **OBSERVATIONS**daten aus gegebenen Pixelpositionen der Blickbewegungsmessung. Die Umsetzung dieser Funktion ist in die Displaysoftware integriert, welche mit der Qt-Bibliothek und Carmenta umgesetzt wurde. Jeder Typ von Anzeigeelement wie beispielsweise ein Textfeld oder ein Knopf wird bei dieser Bibliothek mit eigenen C++-Klassen umgesetzt. Diese Klassen wurden für die semantische Analyse durch eine Analysefunktion erweitert, welche bei Funktionsaufruf ein Objekt zurückliefert, welches das Anzeigeelement eindeutig identifiziert und zusätzlich den Stand der angezeigten Informationen zum Zeitpunkt der Funktion enthält. Bei gegebener Pixelposition wurde ein Mechanismus verwendet, der für die Identifikation von der Maus angeklickter Anzeigeelemente vorgesehen ist, um das angeblickte Element zu identifizieren und die Analysefunktion aufgerufen. Aus dem vom Anzeigeelement zurückgelieferten Datenobjekt wurden dann Observations generiert und über die Interprozesskommunikation versendet. Dieser Mechanismus wurde analog für die Simulationsbestandteile, die nicht auf Qt basieren (beispielsweise X-Plane oder Carmenta) umgesetzt, wobei die Identifikation der angeblickten Elemente sowie die Generierung der Informationen mittels der Schnittstellen dieser Softwarewerkzeuge umgesetzt wurden.

## Editor zur Erstellung des SB-Modells

Zur Erstellung des **SB-MODELLS** wurde das Programm *SitAwEditor* entwickelt (siehe Abbildung 4-5), mithilfe dessen man **SB-KNOTEN** und **RELATIONEN** sowie die assoziierten **OBSERVATIONEN** und der Normierungskonstanten definieren kann. Nach Erstellung des Modells kann mithilfe der Software eine XML-Datei erstellt werden, die von der Applikation zur Berechnung der dynamischen Modellinstanzen eingelesen wird.

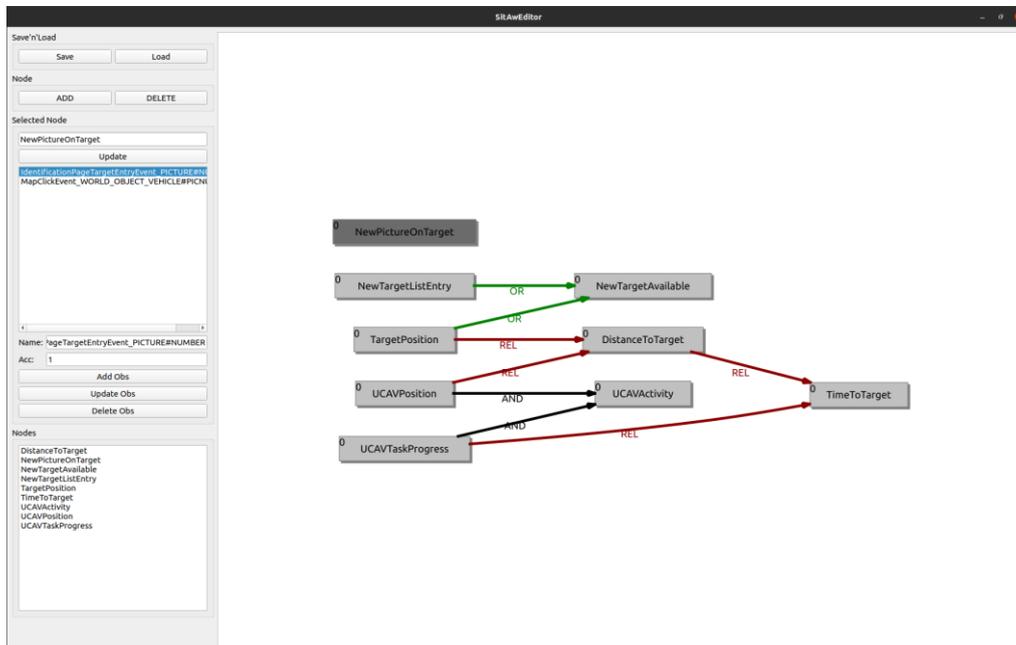


Abbildung 4-5 - Oberfläche des SitAwEditors zur Erstellung von SB-Modellen.

## 4.2 ANALYSE DER VISUELLEN AUFMERKSAMKEIT

Wie bereits ausführlich im Stand der Forschung zur Aufmerksamkeit beschrieben, wird Blickbewegungsmessung in vielen Kontexten zur Analyse der visuellen Aufmerksamkeit eingesetzt. In den genannten Studien wird diese Analyse aber nicht als Klassifikationsproblem betrachtet, sondern mit aggregierten statistischen Maßen durchgeführt (siehe hierzu Abschnitt 2.1.5.). Bei der situativen Betrachtung existieren messtechnische Störfaktoren wie peripheres Sehen oder Unaufmerksamkeitsblindheit, die die Genauigkeit des fixationsbasierten Ansatzes reduziert. Aus diesem Grund steht die Frage nach Richtigkeit und Präzision der vorgeschlagenen fixationsbasierten Analyse im Zentrum der folgenden Studie.

Es soll zunächst geklärt werden, welche messbaren Attribute einer Fixation für die tatsächliche Wahrnehmung von visuellen Informationen sprechen. Hierbei wird ein besonderes Augenmerk auf Fixationsdauer und -abstand zum Interessensgebiet gelegt. Mit Blick auf das vorgestellte Konzept zur Aufmerksamkeitsanalyse ist darüber hinaus auch relevant, inwiefern Interaktion mit dem System und visuelle Aufmerksamkeit auf eine assoziierte Systeminformation zusammenfallen. Zusammenfassend lauten die Forschungsfragen für diesen Versuch:

- Wie valide ist die Klassifikation von Aufmerksamkeit auf Basis von Fixationen im Aufgabenkontext eines Flugzeugcockpits?
- Welche messbaren Attribute einer Fixation sprechen für die Wahrnehmung und Verarbeitung einer Information?
- Können manuelle Interaktion des Bedieners als zusätzliche Indikatoren für visuelle Aufmerksamkeit verwendet werden?

### 4.2.1 GESTALTUNG DER STUDIE

Um die Forschungsfragen im Aufgabenumfeld eines Flugzeugcockpits zu untersuchen ist es notwendig, eine objektive Datenbasis zu schaffen, die mit hoher Genauigkeit klassifiziert, wann eine Information tatsächlich wahrgenommen und verarbeitet wurde. Hierfür wurde ein Studiendesign entwickelt, das sich an Reaktionsexperimenten bei der Untersuchung von Aufmerksamkeit orientiert. Bei der Übertragung auf eine Cockpitumgebung gibt es die folgenden Herausforderungen:

- Die Präsentation von Stimuli ist nicht direkt kontrollierbar, da das Cockpit und die Außenwelt hier grundsätzlich eine Vielzahl von Stimuli bereithalten. Die Unterscheidung, welchem Stimulus aus allen verfügbaren Reizen ein Proband seine Aufmerksamkeit schenkt, ist eine zentrale Frage.
- Ein Versuchsteilnehmer ist nicht passiver Beobachter, der mit Knopfdruck auf einen Reiz reagiert, sondern aktiver Gestalter der Situation, in der er sich befindet. Durch die Aktionen nimmt er Einfluss auf die Versuchsumgebung und Aufgabensituation.
- Die Natürlichkeit der Arbeitssituation geht mit einer Natürlichkeit der Aufgabenausführung einher. Wird beispielsweise die Reaktionszeit auf ein Warnsignal gemessen, ist nicht sicher, dass ein Versuchsteilnehmer stets so rasch wie möglich reagiert, da die sofortige Reaktion im realen Cockpit nur in seltenen Ausnahmefällen notwendig ist.

Mit Blick auf diese Schwierigkeiten wurde für den Versuch das in Abbildung 4-6 dargestellte Konzept umgesetzt:

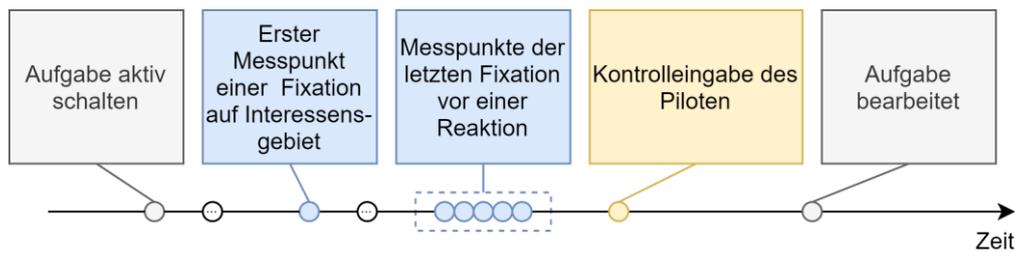


Abbildung 4-6 - Versuchskonzept zur objektiven Messung von Aufmerksamkeit im natürlicher Arbeitsumgebung

Den Versuchsteilnehmern werden mehrere Aufgaben gestellt, die gleichzeitig verfolgt werden sollen. In der Simulation werden diese Aufgaben zu Zeitpunkten aktiviert, die dem Probanden nicht bekannt sind. Aktivieren heißt, dass die Aufgabenumgebung so verändert wird, dass eine Reaktion des Piloten notwendig ist, um seiner Aufgabe nachzukommen. Wenn der Pilot dann reagiert, kann mit großer Sicherheit davon ausgegangen werden, dass der Pilot die aufgabenrelevanten Größen wahrgenommen hat. Dadurch können die Fixationen unmittelbar vor der gemessenen Reaktion analysiert werden, um die eingangs gestellten Forschungsfragen zu evaluieren. Der folgende Abschnitt beschreibt die konkrete Versuchsgestaltung, mit denen dieses Konzept realisiert wurde.

#### 4.2.1.1 Versuchsgestaltung

Der Versuch bestand aus zwei Mission mit unterschiedlichen Aufgabensituationen, für die das Versuchskonzept jeweils umgesetzt wurde. Die erste Mission war ein Instrumentenflug, der durch Windböen gestört wurde. Der zweite Versuch konzentrierte sich auf die Aufgaben der UAV-Führung. In der Auswertung werden die Ergebnisse beider Versuche zusammen vorgestellt.

#### Mission 1: Instrumentenflug mit Anweisungen unter Böeneinfluss

In der ersten Mission mussten die Versuchspersonen kontinuierlich eine vorgegebene Flughöhe, Fluggeschwindigkeit und Route verfolgen. Die Vorgabe war, der Höhe mit einer Toleranz von +200 Fuß, der Geschwindigkeit mit +15 Knoten und der Route mit einem lateralen Abstand von weniger als 2 nautischen Meilen zu folgen.

Voraussetzung für das Versuchskonzept ist die Aktivierung von Aufgaben, um eine Reaktion zu beobachten. Dies wurde für die Höhen- und Geschwindigkeitsaufgabe durch simulierte Böen umgesetzt, die entweder die Höhe oder die Geschwindigkeit störten und in einer Abweichung von 7%-14% über oder unter dem Sollwert resultierten. Die genaue Auswahl ist in dem Abschnitt E.2.1 im Anhang zu entnehmen. Gleichzeitig zu den Böen trafen Nachrichten, die die Zielparameter für Höhe und Geschwindigkeit änderten. Diese Nachrichten mussten über einen Knopf quittiert werden. Die Tabelle 4-2 gibt eine Übersicht über die Nachrichten.

Zeitpunkt	Nachricht
00:00:01	FALCON CLEARED FOR T/O RWY 08 R
00:01:45	FALCON CLB 4500 ft VEL 350 kts
00:07:25	ACCELERATE 390 kts ALT 4500 ft
00:12:55	DECELERATE 360 kts ALT 4500 ft
00:19:20	DECELERATE 340 kts ALT 4500 ft
00:23:52	ACCELERATE 360 kts ALT 4500 ft

Tabelle 4-2 - Übersicht über die Textnachrichten in Mission 1. Vorgabe von Fluggeschwindigkeit (VEL/ACC/DECELERATE) und Flughöhe (ALT).

Bei der Verfolgung der Route gab es die Besonderheit, dass die Route nur bis zum nächsten Wegpunkt auf der Karte sichtbar war und der weitere Verlauf erst bei Ankunft am nächsten Wegpunkt offengelegt wurde, was der Aktivierung der Routenaufgabe entsprach. Der genaue Verlauf der Route bestand aus 10 Wegpunkten, wobei die Flugzeit zwischen zwei Punkten bei Beachtung der Sollgeschwindigkeit zwischen zwei und sechs Minuten lag. Der Routenverlauf ist in Abbildung 4-7 illustriert, in dessen Zentrum der Startpunkt des Flugzeugs liegt.

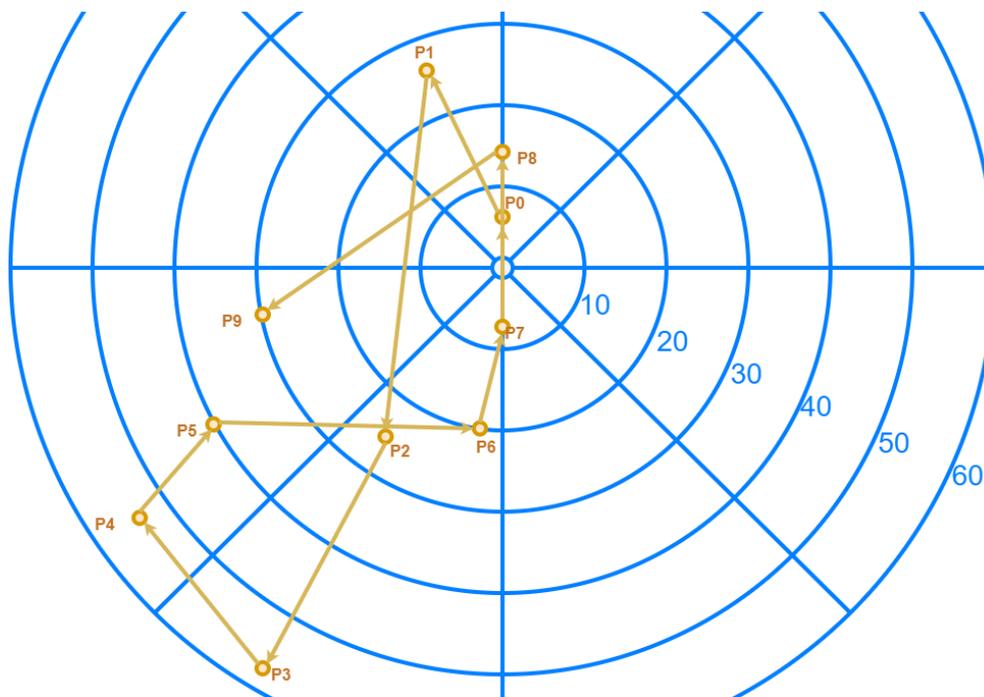


Abbildung 4-7 - Abzufliegende Strecke der Mission 1 bestehend aus zehn Wegpunkten

Eine zusätzliche Aufgabe lag in der Durchführung eines Notfallprozedere in den Zeiträumen von [16:45,17:13] sowie [25:39, 31:05]. Da diese Aufgabe keine nennenswerten Ergebnisse verglichen mit den anderen Zeiträumen des Versuchs vorgebracht hat, wurde der Zeitraum aus der Datenauswertung ausgeschlossen. Die gesamte Durchführungsdauer des Versuchs betrug etwa 30 Minuten.

## Mission 2: Führung von mehreren UAVs

Im Zentrum der zweiten Mission stand die visuelle Aufmerksamkeit auf der taktischen Karte. In diesem Versuch mussten die Probanden drei unbemannte Luftfahrzeuge führen und gleichzeitig ihr eigenes Flugzeug mittels Autopiloten steuern. Abbildung 4-8 zeigt alle Objekte, die im Laufe der Mission für den Probanden sichtbar werden. Der Proband hatte die Aufgabe, mit seinen unbemannten Plattformen alle erscheinenden Gebäudetypen aufzuklären und alle Flugabwehrstellungen zu bekämpfen. Dies konnte er mittels einem Aufgabenzuweisungsdialog direkt auf der taktischen Karte zuweisen (siehe hierzu E.1). Im Versuch erschien durchschnittlich alle 21 Sekunden ein Objekt, wobei zu ausgewählten Zeitpunkten auch zwei gleichzeitig sichtbar wurden. Die Probanden wurden auch aufgefordert, die Ziele möglichst zeitnah zu bearbeiten und ihre Plattformen dafür im Missionsgebiet zu verteilen. Gleichzeitig zu der UAV-Führungsaufgabe erhielt ein Proband zu vorher nicht bekannten Zeitpunkten Anweisungen mit Vorgaben für die Programmierung des Autopiloten. Die Steuerung mithilfe des Autopiloten wurde nicht als eigenständige Aufgabe berücksichtigt und wird als

Folgehändlung der Nachrichtenquittierung gesehen. Der genaue Ablauf der Mission Nachrichten und die Zeitpunkte auftauchender Ziele ist in E.2.2 aufgeführt.

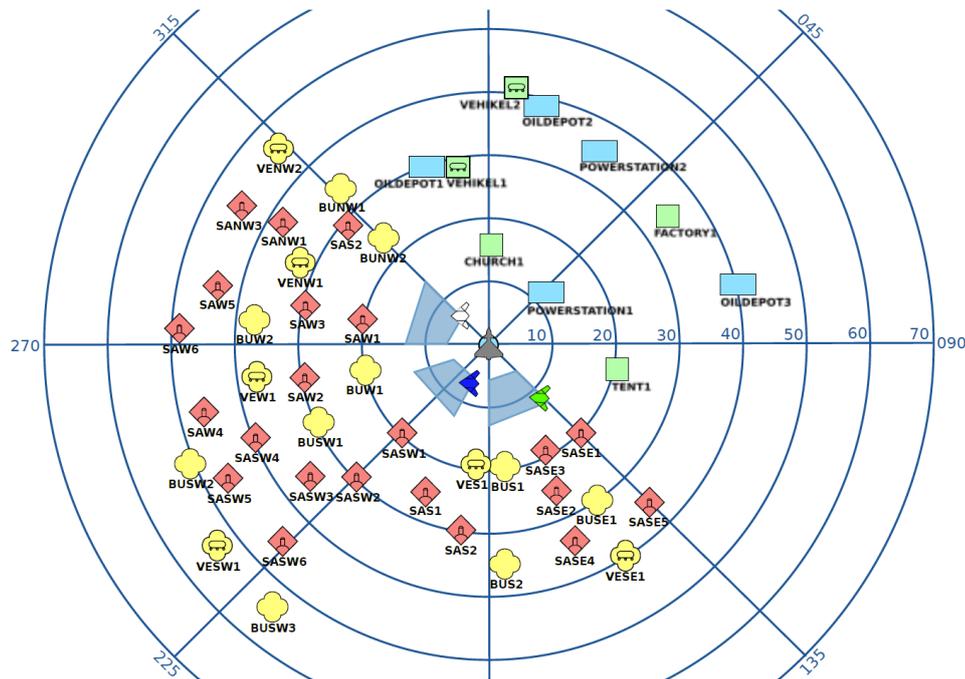


Abbildung 4-8 - Übersicht über die Lage in der Mission 2.

Eine zusätzliche Aufgabe lag in der Durchführung von Notfallprozedere im Zeitraum von [08:50,12:00] sowie [16:57, Ende der Mission]. Da diese Aufgabe keine nennenswerten Ergebnisse verglichen mit den anderen Zeiträumen des Versuchs hervorgebracht haben, wurden die Zeiträume aus der Datenauswertung ausgeschlossen. Die gesamte Durchführungsdauer des Versuchs betrug etwa 25 Minuten.

### Cockpitgestaltung und Interessensgebiete

Abbildung 4-9 zeigt die Gestaltung der Oberfläche mit allen Interessensgebieten. Im HUD sind die Fluggeschwindigkeit und -höhe dargestellt. Dabei sind beide Anzeigen von einer sich bewegenden „Leiter“ umgeben sind, die die aktuelle Veränderung der jeweiligen Größe durch ihre Bewegung visualisiert. Im linken Bildschirm wird im unteren Bereich eine neue Nachricht angezeigt und durch ein Aufleuchten des Labels „DATA COM LINK“ signalisiert. Das Flugzeug und der Flugpfad sind im Interessensgebiet 4 dargestellt und das taktische Objekt im Interessensgebiet 5. Um den Störfaktor der Wahrnehmung durch peripheres Sehen in diesem Versuch zu reduzieren, wurde die Außensicht verdunkelt. Der Proband hatte damit lediglich das HUD und die dort angezeigten Informationen zur Verfügung, um eine Abweichung von den gewünschten Flugparametern festzustellen.

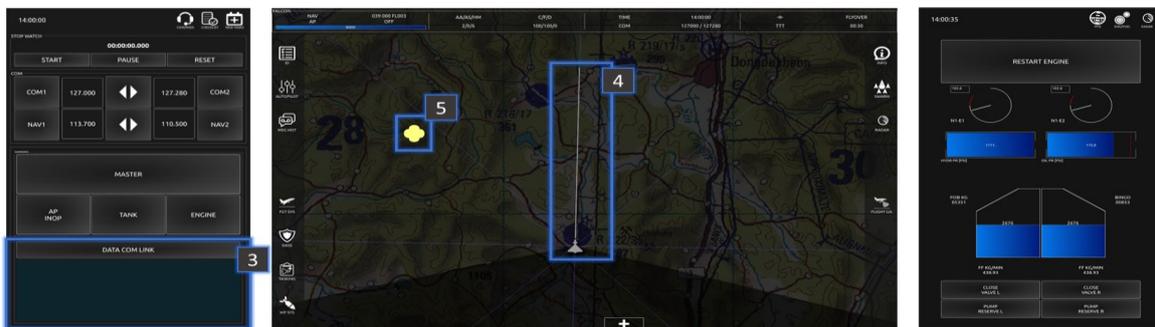
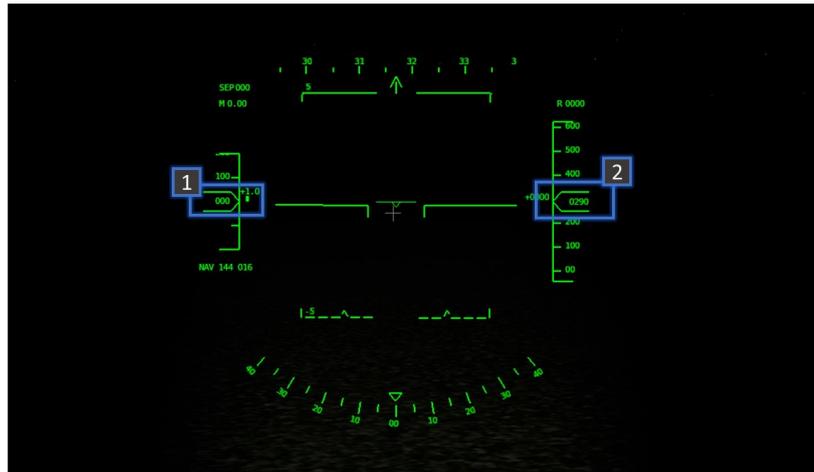


Abbildung 4-9 - Gestaltung der Oberfläche für Mission 1 und 2. Die blau markierten Interessensgebiete sind Flugeschwindigkeit (VEL, 1) Flughöhe (ALT, 2), Nachricht (MSG, 3), Flugpfad (ROUTE, 4) und taktisches Objekt (TO, 5).

#### 4.2.1.2 Messgrößen & Datenverarbeitung

Der Versuchsaufbau erlaubt die Aufzeichnung der simulierten Welt, sowie der Blickbewegungs- und Interaktionsdaten. Die Tabelle 4-3 zeigt, welche Daten bei der Durchführung der Versuche aufgezeichnet wurden.

Messdaten	
<b>Zustand der simulierten Welt</b>	Flugzeugposition, -lage und -geschwindigkeit Position und Sichtbarkeit der UAVs und taktischer Objekte (Gebäude, SAM)
<b>Flug- und Missionsplan</b>	Zustand Routenplan (Wegpunkte) Zustand Missionsplan (UAV-Aufträge)
<b>Steuereingaben</b>	Bewegung der beiden Schubhebel, Eingaben zur Steuerung der Luftbremse Laterale und longitudinale Bewegung des Steuerknüppels Interaktionsposition auf Touchscreen
<b>Blickbewegungsmessung</b>	Position und Zeitstempel des Blicks Klassifikation zwischen Sakkade, Fixation und Blinzeln

Tabelle 4-3 - Aufgenommene Messdaten im Versuch

## Klassifikation der Steuereingaben

Das vorgestellte Versuchskonzept sieht die exakte Messung von Kontrolleingaben in Reaktion auf eine Aufgabe vor. Die Kontrolleingaben wurden wie folgt klassifiziert:

Aufgabentyp	Klassifikation
ALT	Absolute Änderung des Höhenrudersteuers $> 0.075^\circ$ in Richtung der Zielhöhe <i>Oder</i> Bedienung der Höhenrudertrimmung
VEL	Relative Änderung der Schubhebelposition $> 10\%$ in Richtung der Zielgeschwindigkeit <i>Oder</i> Bedienung der Luftbremse
ROUTE	Absolute Änderung des Querrudersteuers $> 0.04^\circ$ in Richtung des Zielkurses
MSG	Bedienung der Nachrichtenquittierung
UAV	Anklicken des Ziels zum Aufrufen des Kontextmenüs

Tabelle 4-4 - Klassifikationsregeln für Reaktionen auf eine Aufgabe

Die bereits beschriebene Natürlichkeit der Versuchsumgebung erschwert eine exakte Klassifizierung vor allem für die kontinuierlichen Eingaben. Die Grenzwerte hierfür wurden durch Analyse und Visualisierung der Versuchsdaten festgelegt. Die Abbildung 4-10 zeigt hierfür beispielhafte Daten für die Aufgabe ALT.

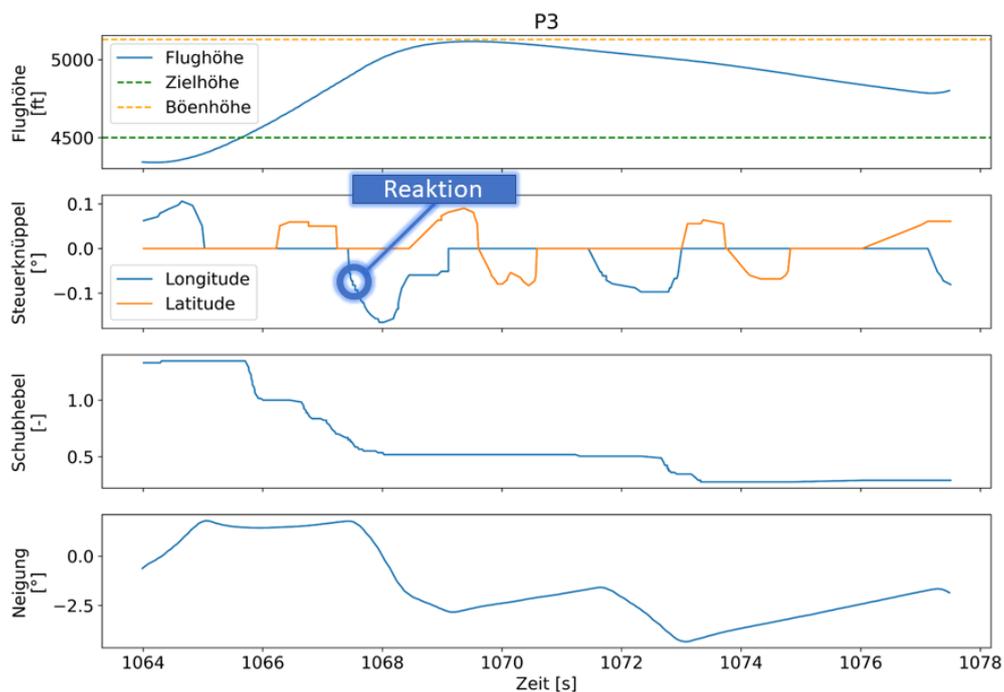


Abbildung 4-10 - Verlauf von Flughöhe, Steuerknüppelausschlag, Schubhebelstellung und Flugzeugneigung nach einer Windböe bei  $T=1064s$  mit markierter Reaktion in der Steuerknüppelstellung ist markiert.

## Datenverarbeitung

Zur Bewertung der Validität werden im Folgenden Metriken zur Bewertung von Klassifikationssystemen herangezogen. Hierfür werden Prinzipien der Signalentdeckungstheorie

angewendet, welche sich mit der Analyse von Treffer- und Fehlerrate in Abhängigkeit von möglichen Grenzwertparametern in Erkennungssystemen beschäftigt (Metz, 1978). Die dort entwickelten Metriken beschreiben das Verhältnis der getroffenen Klassifikationen zu den Referenzklassen einer Messung. Hier kommen die folgenden Metriken zum Einsatz:

$$\text{Präzision} = \frac{n_{\text{richtig pos}}}{n_{\text{richtig pos}} + n_{\text{falsch pos}}} \quad (4)$$

$$\text{Sensitivität} = \frac{n_{\text{richtig pos}}}{n_{\text{richtig pos}} + n_{\text{falsch neg}}} \quad (5)$$

Ferner werden die folgenden Variablen bei der Auswertung verwendet:

$$\text{Dauer zwischen Fixation und Reaktion} \quad T_{\text{fix-react}}^{(AoI)} = T_{\text{react}}^{(AoI)} - T_{\text{fixation}}^{(AoI)} \quad (6)$$

$$\text{Dauer für Bearbeitung der Aufgabe} \quad T_{\text{task}}^{(AoI)} = T_{\text{task done}}^{(AoI)} - T_{\text{task start}}^{(AoI)} \quad (7)$$

Mittels der Größe  $T_{\text{fix-react}}^{(AoI)}$  werden falsch negative Ergebnisse identifiziert. Eine  $T_{\text{fix-react}}^{(AoI)} < 0$  wird als falsch negative Klassifikation des Messsystems gewertet, da eine Reaktion vor der ersten Fixation gemessen wurde. Ferner wird ein Zeitraum auch als falsch negativ bezeichnet, wenn im Verarbeitungszeitraum gar keine Fixation gemessen wird. Als falsch positiv wurde klassifiziert, wenn  $T_{\text{fix-react}}^{(AoI)}$  größer ist als das obere Quartil ( $p > .75$ ) aller  $T_{\text{task}}^{(AoI)}$  in einer Aufgabe. Diese Grenze ist großzügig gewählt, um der Natürlichkeit der Aufgabe Rechnung zu tragen und auch langsamere Reaktionen als richtig positive Fälle zu klassifizieren.

Statistische Tests wurden mit dem sogenannten Bootstrapping-Verfahren durchgeführt (Efron & Tibshirani, 1993). Vorteil dieser Methode ist, dass keine Annahmen über die Verteilung der Teststatistik in der Population getroffen werden muss, sondern diese Verteilung aus den Daten abgeschätzt wird. Ferner können beliebige Teststatistiken getestet werden. Das Konfidenzintervall wurde zu  $p = 0.05$  gewählt.

Datenauswertung sowie Durchführung der statistischen Tests wurde mit Python und den Bibliotheken *Scipy*, *Numpy* sowie *Pandas* durchgeführt<sup>11</sup>.

<sup>11</sup> Website: Scipy (<https://scipy.org/>), Numpy (<https://numpy.org/>), Pandas (<https://pandas.pydata.org/>)

#### 4.2.1.3 Teilnehmer

Die Studie wurde mit zehn männlichen Probanden durchgeführt. Das Durchschnittsalter der Probanden beträgt 36.9 Jahre. Alle Probanden waren zum Zeitpunkt der Versuche im Besitz einer Pilotenlizenz. Davon waren fünf Probanden Halter einer Verkehrspilotenlizenz (ATPL). Die niedrigwertigste Lizenz war eine Segelflugglizenz mit Klassenberechtigung Motorsegler (SPL mit TMG).

ID	Alter	Berechtigungen	Typ	Flugstunden
1	23	2	Privatpilot	500
2	35	2	Privatpilot	550
3	59	3	Berufspilot	20000
4	32	3	Berufspilot	1367
5	24	1	Privatpilot	350
6	28	2	Privatpilot	582
7	24	1	Privatpilot	400
8	55	3	Berufspilot	29170
9	46	3	Berufspilot	8380
10	29	3	Berufspilot	981

Tabelle 4-5 - Versuchsteilnehmer. Berechtigungen: 0: Glider; 1: TMG; 2: SEP/MEP Single Pilot; 3: MEP MultiCrew

Aufgrund von Problemen in der Datenaufzeichnung zu Beginn der Studie musste ein Proband (ID=1) aus der Studie ausgeschlossen werden.

#### 4.2.1.4 Ablauf

Den Probanden wurde vorab ein Simulator- und Missionshandbuch zugesandt (Hoffelner, 2021). Danach erhielten alle Probanden ein Training von annähernd einer Stunde im ersten Versuchsszenario. Während dieses Trainings begegnete den Probanden alle Versuchsaufgaben der ersten Mission. Dann wurde das Blickbewegungssystem kalibriert mit einer durchschnittlichen Genauigkeit von  $1.04^\circ$  ( $SD = 0.72^\circ$ ). Dann führten die Probanden die erste Mission durch. Im Anschluss wurde ein zweites Training durchgeführt, das mit dem zweiten Versuchsszenario vergleichbar war und alle Aufgaben enthielt. Zuletzt wurde der zweite Versuch durchgeführt.

### 4.2.2 ERGEBNISSE

Für die Auswertung der Ergebnisse wurden die Messdaten beider Missionen kombiniert. Aus der ersten Mission werden die vier Aufgaben ROUTE, ALT, VEL und MSG betrachtet und aus der zweiten Mission nur die Aufgaben TO (*Tactical Object*). Es stellte sich nach den Versuchen heraus, dass eine neue Nachricht zur Programmierung des Autopiloten nicht zur unmittelbaren Quittierung der Nachricht führte, sondern einige Piloten erst den Autopiloten bedienten, bevor sie die Nachricht quittierten. Aus diesem Grund wurden diese Daten aus der Auswertung ausgeschlossen. Die folgenden Ergebnisse wurden aus den Blickbewegungsdaten aller Piloten erstellt. Der betrachtete Zeitraum begrenzt sich ausschließlich auf den Zeitraum zwischen Aktivschalten einer Aufgabe und Bearbeitung der Aufgabe liegen (siehe Abbildung 4-6). Eine Ausnahme hiervon ist der Vergleich der Fixationsdauern, bei dem einmal alle Fixationen ausgewertet wurden.

#### Richtigkeit & Präzision

Der erste Blick auf die Ergebnisse zielt auf eine Einschätzung von Richtigkeit und Präzision ab. Hierbei ist zunächst zu klären, was als Klassifikationsfehler gewertet wird. Hierfür zeigen die Diagramme in Abbildung 4-11 die Verteilung von allen  $T_{fix-react}$  getrennt nach Aufgabentypen über variierte Fixationsradien. Dabei werden negative  $T_{fix-react}$  abgebildet, welche dann auftreten, wenn eine

Reaktion des Probanden vor einer ersten Fixation auf das Interessensgebiet gemessen wird. Nicht abgebildet sind die Datenpunkte, bei denen keine Fixation auf das Interessensgebiet gemessen wurde, bevor eine Aktion durchgeführt wurde. Beide Fälle – fehlende Fixation und negative  $T_{fix-react}$  – sind falsch negative Ergebnisse.

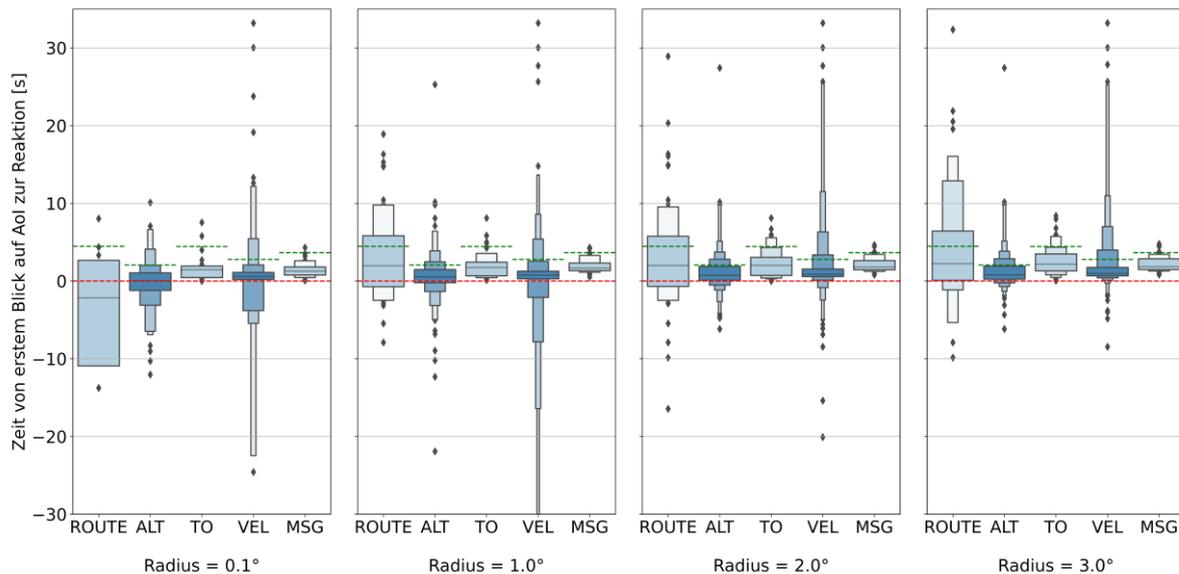


Abbildung 4-11 - Zeit  $T_{fix-react}$  zwischen Beginn der ersten Fixation und Reaktion in Abhängigkeit von Aufgabe und Auswertungsradius. Referenz auf null (rot-gestrichelt) und Referenz auf das obere Quartil der Reaktionszeit  $T_{task}^{(Aoi)}$  (grün-gestrichelt).

Alle Datenpunkte, die zwischen der Nullreferenz und dem oberen Quartil von  $T_{task}^{(Aoi)}$  (grün-gestrichelt) liegen, werden als richtig positive Klassifikation gewertet. Das vorgestellte Versuchsdesign kann keine objektive Aussage über richtig negative Klassifikationen geben, da in diesen Fällen keine messbare Reaktion des Bedieners vorhanden ist. Auffällig ist, dass ein kleiner Auswertungsradius von  $r = .1^\circ$  bei der Routenaufgabe zu fast ausschließlich falsch negativen Ergebnissen führt. Dem gegenüber steht die Nachrichtenaufgabe, deren Klassifikationsergebnisse bei Änderung des Radius kaum ändern. Dies bedeutet, dass hier eine richtige Klassifikation auch bei kleinem Auswertungsradius stattgefunden hat.

Bei den taktischen Objekten (TO) gab es den Fehler in der Versuchsgestaltung, dass mehrere Objekte gleichzeitig auftauchten. Dies entspricht nicht dem Versuchskonzept, da hier die Wahrnehmung eines Objekts nicht mit einer instantanen Reaktion einhergeht, sondern mehrere Ziele nacheinander abgearbeitet werden. Um dieses Problem zu beheben, wurde nur das zuerst bearbeitete der gleichzeitig auftauchenden Objekte ausgewertet.

Appendix Tabelle 4 im Anhang E zeigt die genaue Anzahl von falschen und richtigen Klassifikationen, welche in Abbildung 4-12 visualisiert werden. Mit steigendem Auswertungsradius sinkt die Anzahl der falsch negativen Ergebnisse, während die Zahl der richtig positiven stark ansteigt. Dies wird durch die stark steigende Sensitivität abgebildet. Dem gegenüber steigt die Anzahl der falsch positiven Ergebnisse mit steigendem Radius, was sich in einer sinkenden Präzision niederschlägt. Zwischen  $r = 2^\circ$  und  $r = 3^\circ$  sinkt die Präzisionsrate unter die Sensitivitätsrate, was bedeutet, dass ein großer Auswertungsradius zu einer erhöhten falsch-positiven Klassifikationsrate führt. Bei Auswertungsradien  $r > 3^\circ$  ändern sich die Ergebnisse nur noch marginal.

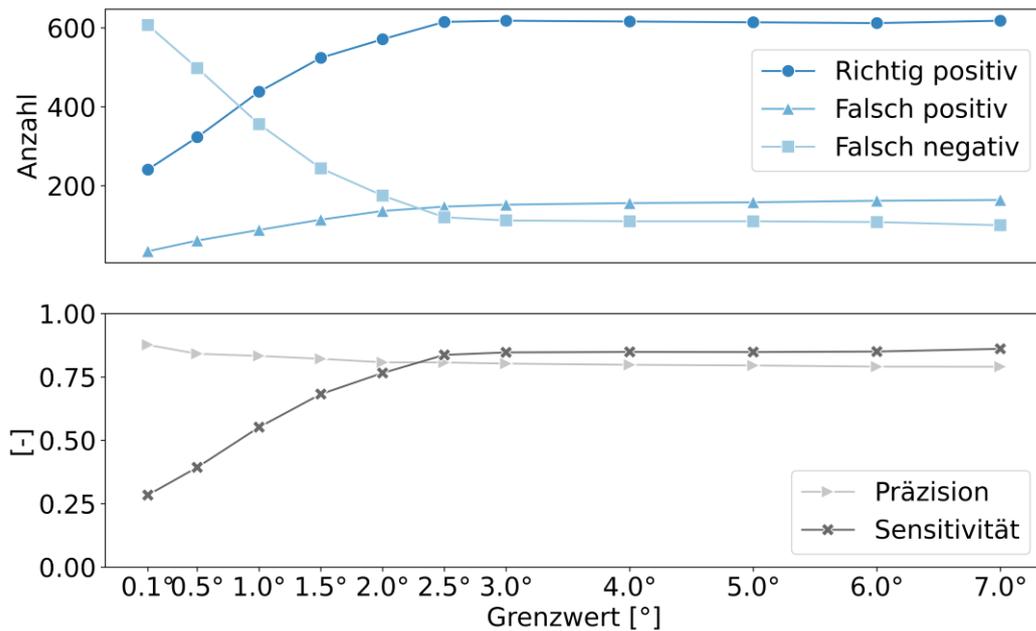


Abbildung 4-12 - Oben: Verhältnis von Fehler- zu Trefferrate in Abhängigkeit vom Auswertungsradius. Unten: Präzision und Sensitivität im Verhältnis zum Auswertungsradius.

### Fixationsdauer

Das Ergebnis der Analyse der Fixationsdauer ist in Tabelle 4-6 aufgeführt und die Verteilung der Messdaten ist in Abbildung 4-13 dargestellt.

Fixation	N	Durchschnitt [s]	Median [s]	Std [-]
Alle	62523	0.330	0.220	0.408
Vor Aktion	1045	0.401	0.260	0.412
ROUTE vor Aktion	67	0.303	0.160	0.374
ALT vor Aktion	263	0.314	0.220	0.322
TO vor Aktion	247	0.389	0.236	0.442
VEL vor Aktion	415	0.466	0.361	0.411
MSG vor Aktion	53	0.599	0.660	0.560

Tabelle 4-6 - Vergleich des Durchschnitts, Medians und der Standardabweichung (Std) der Fixationsdauern

Zunächst kann festgestellt werden, dass die durchschnittliche Fixationsdauer aller Fixationen mit  $\bar{T}_{fix} = 330ms$  im Bereich der typischen Werte aus der Literatur liegt, wobei auch ein großer Anteil von kurzen Fixationen mit Dauern zwischen 50 – 330ms gemessen wurden. Bei derart kurzen Fixationen ist fraglich, ob die relevanten Informationen bereits in vorhergehenden Fixationen wahrgenommen wurden. Der Durchschnitt der Fixationsdauer ( $\bar{T}_{fix,vor\ Aktion} = 401ms$ ) kurz vor einer Reaktion liegt etwas höher als bei allen gemessenen Fixationen, wobei sich die Verteilung der Dauern stark überlappt und damit kein robuster Grenzwert zur Unterscheidung existiert.

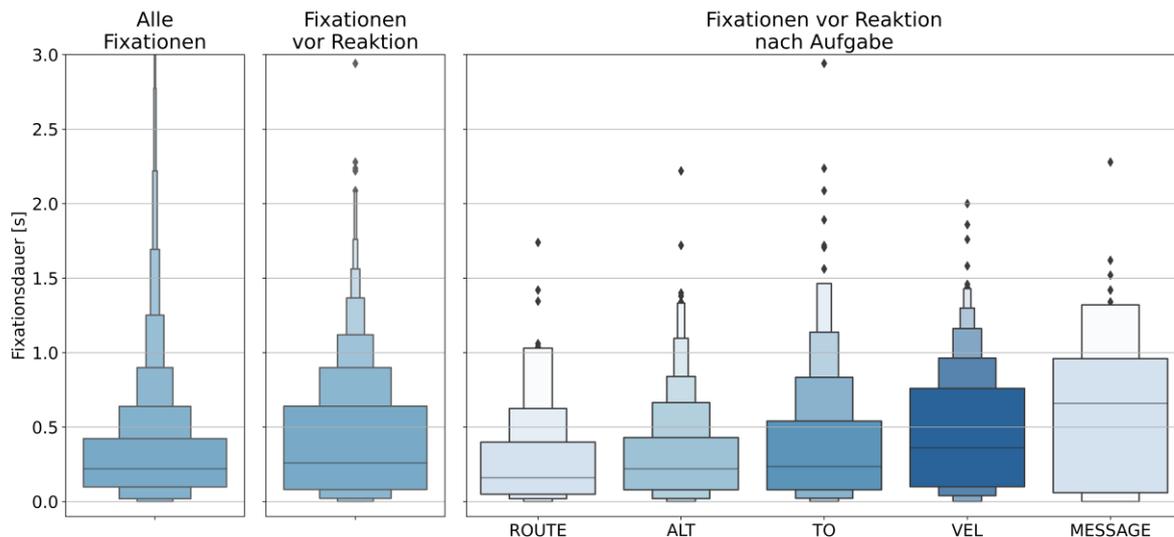


Abbildung 4-13 - Vergleich aller Fixationsdauern mit den Fixationsdauern vor Reaktion auf die Aufgaben ROUTE, ALT, TO, VEL und MESSAGE.

Einen interessanten Aspekt zeigt der Vergleich der Fixationsdauern zwischen den Aufgabentypen. Das Interessensgebiet für die Routenaufgabe hat mit einer  $\bar{T}_{fix,ROUTE} = 303ms$  eine wesentlich niedrigere Fixationsdauer als die für Geschwindigkeit ( $T_{fix,VEL} = 466ms$ ) und Textnachricht ( $\bar{T}_{fix,MSG} = 599ms$ ). Ein plausibler Grund hierfür wäre, dass die Interessensgebiete mit niedrigeren Fixationsdauern bereits vor der direkten Fixation im peripheren Sichtfeld verarbeitet werden oder dass die Informationsaufnahme dort schneller ist. Das könnte auch erklären, warum das Lesen der Textnachricht längere Fixationen benötigt als die Wahrnehmung einer Bewegung der Höhenleiter im HUD.

### Interaktion und Blickbewegung

Um zu untersuchen, wie Interaktion und Blickbewegung korrelieren wurde nicht wie bisher die Fixation vor der Aktion betrachtet, sondern die Fixation zum Zeitpunkt der Aktion. Das Histogramm in Abbildung 4-14 zeigt die gemessene Winkeldistanz dieser Fixation zum Zentrum des aufgabenbezogenen Interessensgebiets. Es fällt auf, dass bei den Interessensgebieten ROUTE, TO und MSG der größte Anteil der Fixationen sehr nah am Interessensgebiet liegt. Dem gegenüber stehen die Blickdistanzen im HUD mit einer breiteren Verteilung bei VEL und ALT. In beiden Interessensgebieten sind drei Distanzanhäufungen bei etwa  $0^\circ$ ,  $10^\circ$  und  $20^\circ$  Blickdistanz zu sehen. Der Abstand von  $20^\circ$  entspricht dem Interessensgebiet von VEL beziehungsweise ALT. Die mittlere Distanz liegt im Zentrum des HUDs. Es zeigt sich also, dass die Versuchspersonen oft schon auf eine Böe reagierten, ohne direkt auf das Interessensgebiet zu blicken. Dieses Verhalten ist bei ALT noch ausgeprägter als bei VEL.

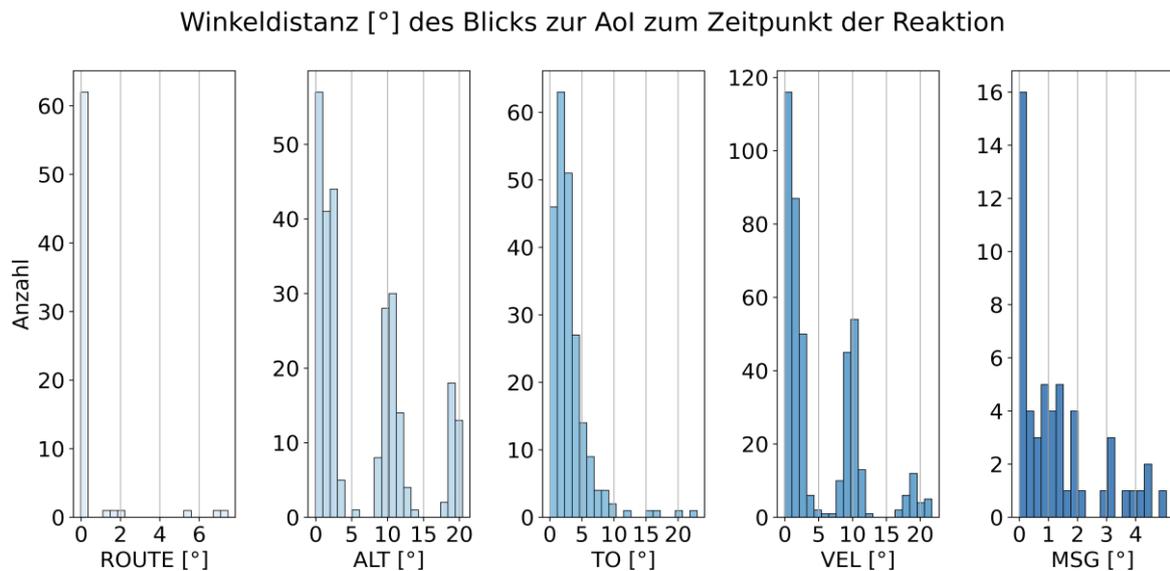


Abbildung 4-14 - Winkeldistanz des Blicks zum aufgabenbezogenen Interessensgebiet zum Zeitpunkt einer Aktion

### 4.2.3 DISKUSSION

Die Ergebnisse der beiden Missionen zeigen im Zusammenhang mit der Forschungsfrage nach Validität, dass eine fixationsbasierte Messung von visueller Aufmerksamkeit valide Ergebnisse erzielen kann und mit hoher Wahrscheinlichkeit ein Großteil der Fixationen im Cockpit der Wahrnehmung von Informationen gilt. Wie die Abbildung 4-11 zeigt, gehen dem Großteil der Reaktionen Fixationen voraus. Dieses Ergebnis ist deckungsgleich mit der Forschungsliteratur und rechtfertigt auch die verbreitete Verwendung von fixationsbasierten Ansätzen. Der Versuch deckt aber auch Probleme des Messverfahrens auf, die zu Messfehlern in der weiteren Verwendung der Ergebnisse führen kann. Beispielsweise zeigen die negativen Zeiten in Abbildung 4-11, dass es auch eine hohe Anzahl an falsch negativen Klassifikationen gibt, die sich auch in der Sensitivität niederschlägt.

Die zweite Forschungsfrage zielt auf die Gestaltung einer fixationsbasierten Klassifikation. Der Versuch konnte eine objektive Basis zur Beantwortung dieser Frage bereitstellen und es zeigt sich, dass die exakte Auswertung des Fixationspunkts eine hohe Präzision erreicht, aber gleichzeitig eine hohe Rate an falsch negativen Ergebnissen liefert. Deshalb ist die Einführung eines Auswertungsradius notwendig, denn es zeigt sich, dass bei ansteigenden Auswertungsradien die hohe Fehlerrate reduziert werden kann. Bei zu großem Auswertungsradius steigt jedoch der Fehler der falsch positiven Klassifikationen. Überraschend ist, dass sich die Präzision und Sensitivität im Versuch selbst bei großen Auswertungsradien nicht mehr stark änderte (vgl. Abbildung 4-12). Ein möglicher Grund hierfür ist, dass die Interessensgebiete im ersten Versuch räumlich stark getrennt waren und somit selbst große Radien keine nennenswerte Änderung der Klassifikationen in Richtung falsch positiver Einschätzung verursachten. Anders ist das beim zweiten Versuch, bei dem die taktischen Objekte durchaus eng angeordnet waren. Hier gab es den Störfaktor, dass alle gleichzeitig auftauchenden Objekte bemerkt wurden, die durchzuführende Delegation von Aufträgen aber nacheinander durchgeführt wurde und somit kein valides Ergebnis mehr im Sinne der Versuchsgestaltung erreicht werden konnte. Hierfür wurden, wie bereits beschrieben, die Daten so gefiltert, dass nur das zuerst bediente taktische Objekt ausgewertet wurde. Aber selbst hier zeigte sich eine lange Zeit zwischen Auftauchen der Objekte und Reaktion der Bediener.

Im Versuch konnte ein Einfluss der Gestaltung von Information auf die Fixationsdauern nachgewiesen werden (siehe Abbildung 4-13). Der Unterschied zwischen der Text- und der Routenaufgabe ist besonders prägnant. Es ist plausibel, dass dies an unterschiedlicher Verarbeitung der Text- gegenüber der Kartendarstellung liegt. Nicht gesichert lässt sich sagen, ob dies an peripherer Wahrnehmung oder an schnellerer Aufnahme von grafischen Informationen gegenüber Text liegt. Dass nicht nur die Art der Darstellung, sondern die Verarbeitung von Bewegungen im peripheren Sichtfeld eine Rolle spielt, zeigt sich bei der Aufgabe ALT und VEL. Zunächst ist die Fixationszeit verglichen mit den anderen Aufgaben kurz, aber darüber hinaus liegt der Blickpunkt zum Zeitpunkt der Reaktion oft gar nicht direkt beim Interessensgebiet (siehe Abbildung 4-14). Dies lässt sich dadurch erklären, dass die Bewegung der Höhen- und Geschwindigkeitsleiter im peripheren Blickfeld wahrgenommen wird und sofort zu einer Reaktion führt. Dies ist ausgeprägter bei der Aufgabe ALT, da die Bewegung der Höhenleiter verglichen mit VEL wesentlich schneller ist.

Abschließend sollte der Zusammenhang zwischen visueller Aufmerksamkeit zum Zeitpunkt einer zielgerichteten Interaktion analysiert werden. Dass eine Reaktion des Bedieners zunächst Aufmerksamkeit voraussetzt, war eine Grundannahme der Versuchsgestaltung. Die Ergebnisse des Versuchs zeigten aber, dass der Blickpunkt nicht zwingend zum Zeitpunkt der Reaktion auf den damit verbundenen Informationen liegt. Dies ist vor allem im HUD der Fall, bei dem der Blick zum Zeitpunkt der Reaktion oft an anderer Stelle lag. Dagegen war die Blickdistanz bei den Interessensgebieten auf den Touchscreens (ROUTE, TO und MSG) nah beim Interessensgebiet. Bei Interaktionen auf dem Touchscreen – wie bei TO und MSG – ist dieser Zusammenhang schon im Kontext von Blickbewegungskalibrierung nachgewiesen worden, da der Bediener zur Positionierung des Fingers die visuelle Aufmerksamkeit auf den Punkt der Berührung richten muss (X. Zhang et al., 2018).

### Limitierung der Versuchsgestaltung

Der hier vorgestellte Versuch hat einige Schwachstellen, die Einfluss auf die Validität der Ergebnisse haben. Zunächst konnten keine fehlerfreien Regeln zur Aktionsklassifizierung der kontinuierlichen Steuereingaben mit dem HOTAS-System identifiziert werden. Die verwendeten Regeln (siehe Tabelle 4-4) generieren falsch-negative und falsch-positive Klassifikationen von Aktionen, das heißt Reaktionen können nicht erkannt oder Interaktionen werden fälschlicherweise als Reaktion gewertet. Aufgrund dieser Problematik modellieren Honecker und Schulte (2017) in ihrer Aktivitätserkennung Fehlermodelle für die unterschiedlichen Interaktionen, wobei die Bedienung des Touchscreens wesentlich weniger anfällig für Fehler ist als die des Steuerknüppels. Das entspricht auch den Ergebnissen dieses Versuchs, bei denen diskrete Aktionen wie Knopfdrücke oder Bedienung des Touchscreens fehlerfrei klassifiziert werden können gegenüber den kontinuierlichen Eingangsgrößen bei Steuerknüppel und Schubhebel. Fehlerhafte Klassifikation der Reaktion beeinflussen möglicherweise den Zeitraum, der für die Auswertung herangezogen wird.

Die Probandengruppe bestand aus erfahrenen Berufspiloten und unerfahrenen Privatpiloten. Es wurde in einigen Studien gezeigt, dass die Erfahrung einen Einfluss auf die Blickbewegung wie beispielsweise Fixationsdauern und -frequenzen hat (Lounis et al., 2021). Da das Forschungscockpit kein konventionelles Cockpit abbildet und auch die Aufgabentypen nicht genau denen eines zivilen Piloten entsprechen, könnte diesen Störfaktor etwas abgeschwächt haben.

Die Versuchsgestaltung kann darüber hinaus keine Aussagen über richtig negative Klassifikationen tätigen, da der Zeitraum der Auswertung sich immer auf die Reaktion des Piloten und damit auf eine positive Klassifikation bezieht.

## Zusammenfassung

Aus diesen Ergebnissen können folgende Erkenntnisse für das Messsystem für visuelle Aufmerksamkeit gewonnen werden:

- Die Wahl des Auswertungsradius  $r = 2^\circ$  ist aus drei Gründen eine sinnvolle Wahl. Zunächst ist die Wahl eines kleinen Auswertungsradius sinnvoll, um zwischen eng angeordneten Interessensgebieten unterscheiden zu können. Zusätzlich zeigte sich im Versuch im Bereich um  $r = 2^\circ$  eine gute Präzision und Sensitivität und nicht zuletzt entspricht diese Größe auch den Ergebnissen vergangener Studien über das menschliche Sichtfeld (Duchowski, 2017).
- Die Fixationsdauer als Klassifikationsregel ist nicht sinnvoll, da die Unterschiede nicht groß sind und sich die Verteilung der Fixationen vor Reaktionen stark mit Fixationen ohne Reaktionen überschneiden.
- Vor allem Interaktionen auf den Touchscreens können als Ergänzung von Blickbewegungsmessung genutzt werden. Zum einen war der Zusammenhang im Versuch dort wesentlich stärker ausgeprägt und zum anderen sind diskrete Interaktion dort genauer zu messen als die kontinuierliche Interaktion mit Schubhebel und Steuerknüppel.

### 4.3 BEWERTUNG DES SITUATIONSBEWUSSTSEINS

In der vorhergehenden Studie wurde gezeigt, dass die Analyse von Aufmerksamkeit mit einer ausreichenden Validität und Sensitivität funktioniert. Der folgende Versuch baut darauf auf und soll die Situationsbewusstseinsbewertung untersuchen und damit die Frage, inwiefern die Berechnung einer informationspezifischen Abweichung durch den Vergleich von Bediener- und Systemmodell aussagekräftig für das tatsächliche SB des Bedieners und die damit verbundene Leistungsfähigkeit ist. Ein weiterer Aspekt der folgenden Studie war die Erkennung von Bedienerfehlern aufgrund falschen oder fehlenden SBs mit Hinblick auf die Verwendung in einem adaptiven Assistenzsystem. Mithilfe der Versuche sollen die folgenden Fragen beantwortet werden:

- Besteht ein Zusammenhang zwischen Referenzmaßen von SB, der Leistung und der blickbewegungs-basierten SB-Abweichung eines Bedieners?
- Eignet sich die SB-Bewertung zur situativen Identifikation von Fehlern eines Bedieners?

Das Messsystem, das im Folgenden Einsatz kommt, baut auf der Aufmerksamkeitsmessung und den im vorherigen Versuch gewonnenen Erkenntnissen auf (siehe Abbildung 3-12). Beispielsweise wurde in der Auswertung der optimale Auswertungsradius verwendet und Interaktionen auf den Touchscreens als Ergänzung der Blickbewegungsmessung verwendet.

#### 4.3.1 GESTALTUNG DER STUDIE

Um das Messkonzept zu validieren sind Vergleichsgrößen notwendig, welche Situationsbewusstsein unabhängig von der Blickbewegungs- und Interaktionsmessung messen. Wie in Abschnitt 2.2.2 diskutiert, basieren die etablierten Messverfahren von Situationsbewusstsein auf Abfrage von Informationen während oder bei gestoppter Simulation der Aufgabensituation (bspw. SART oder SAGAT). Die Verwendung dieser Methoden ist problematisch, da sie im Falle der Befragung bei gestoppter Simulation eine unnatürliche Aufgabenunterbrechung benötigen und im Falle der Befragung bei laufender Simulation für zusätzliche Arbeitsbelastung sorgen. Aus diesem Grund wurde bei der Gestaltung der Versuche darauf geachtet, eine Aufgabensituation herzustellen, in der ein direkter Zusammenhang zwischen SB mit Fehlermaßen und Reaktionszeiten besteht. Die zugrundeliegende Annahme, dass ein gutes SB sich direkt in Fehlern beziehungsweise schnellerer Entscheidung niederschlägt, ist nicht unstrittig. Endsley (2019) weist zwar in einer Meta-Analyse auf eine positive Korrelation Leistungsmaßen mit den Ergebnissen von SAGAT und SPAM hin (mit Korrelationskoeffizienten von  $r = [0.41, 0.53]$ ). Die Autoren einer anderen Metastudie kritisieren dabei aber, dass die getroffene Auswahl der Arbeiten hier eine Überschätzung der Korrelation begünstigt und die Korrelation nach einer weniger strengen Auswahl nur sehr schwach und individuell stark schwankend sei (Bakdash et al., 2022). Aus diesem Grund wurde bei der Gestaltung der Versuchsaufgaben darauf geachtet, einen direkten Zusammenhang zwischen Situationsbewusstsein und der Leistung des Bedieners herzustellen. Das heißt, dass ein Bediener mit hohem Situationsbewusstsein eine höhere Leistung im Sinne einer Entscheidungsgeschwindigkeit oder -qualität vorweisen sollte. Dabei wurden die Aufgaben möglichst simpel gestaltet, damit die relevanten Informationen in einem **SB-MODELL** niedriger Komplexität abbildbar sind. Die Ziele des Probanden waren dabei direkt durch die Aufgabenstellung vorgegeben, was die Erstellung eines Modells vereinfachte.

#### 4.3.1.1 Versuchsgestaltung

Die Probanden führten zwei Missionen mit unterschiedlichem Aufgabenprofil durch. Im ersten Versuch musste primär eine anspruchsvolle Tiefflugaufgabe durchgeführt und gleichzeitig unbekannte Bedrohungen identifiziert werden. Die Tiefflugaufgabe entspricht einer sensor-motorischen Aufgabe mit hoher Anforderung an die visuelle Aufmerksamkeit des Bediener, während die Identifikation der Bedrohung eine Entdeckungsaufgabe darstellt. Im zweiten Versuch mussten mehrere UAVs delegiert und Aufklärungsbilder klassifiziert werden. Diese Aufgabe kann als Aufgabe der überwachenden Kontrolle gesehen werden.

##### Mission 1: Tiefflug und Meldung von Flugabwehrstellungen

Die Aufgabe des Versuchsteilnehmers war ein Tiefflug entlang einer Route durch Täler der Alpen. Die Route wurde durch Wegpunkte gekennzeichnet, die der Versuchsteilnehmer überfliegen musste. Die Wegpunkte waren nummeriert und dienten als Anhaltspunkt der aktuellen Position eines Versuchsteilnehmers. In der taktischen Karte war auch das Relief des zu durchfliegenden Bereichs dargestellt, dessen genaue Routenführung in Abbildung 4-15 dargestellt ist. Beim Tiefflug sollte eine Radarhöhe von 500 Fuß nicht überfliegen werden und die Fluggeschwindigkeit zwischen 500 und 600 Knoten gehalten werden. Eine Schwierigkeit war dabei, dass sich das Terrain kontinuierlich veränderte und der Proband gleichzeitig durch enge Täler flog.



Abbildung 4-15 - Flugpfad definiert durch blaue Wegpunkte und Platzierung der auftauchenden Flugabwehrstellungen in Missionsszenario 1

An zehn Positionen entlang der Route befanden sich Flugabwehrraketenstellungen (engl. *Surface-Air-Missile-Site*, SAM-Site) mit der Bezeichnung „SA-14“. Diese waren zunächst nicht auf der Karte sichtbar, sondern erschienen erst, sobald das Flugzeug in einen Radius um die Stellung herum vorgedrungen war. Die Aufgabe der Teilnehmer war es, eine Meldung abzugeben, sobald sie die Flugabwehrstellung bemerkten. Die Meldung sollte die Art der Stellung, den nächsten zu überfliegenden Wegpunkt und die Uhrzeit beinhalten. Diese Reihenfolge war dabei einzuhalten. Die Uhrzeit konnte auf dem mittleren Bildschirm oberhalb der Karte abgelesen werden. Das Szenario war beendet, sobald der letzte Wegpunkt überflogen wurde. Die Probanden erhielten keinerlei auditive Unterstützung. Der Versuch dauerte etwa 16 Minuten.

## Mission 2: Führung von UAVs unter Zeitdruck

Während der zweiten Mission mussten die Probanden drei UAVs zur Aufklärung und Identifizierung von Zielen einsetzen. Das Kampfflugzeug musste nicht geflogen werden, daher erfolgte die Steuerung dabei nur über das MHDD des Cockpits. Im Versuch erschien 22 zu identifizierende Objekte zeitlich verzögert und in den Probanden unbekannter Reihenfolge auf der Karte. Die Aufgabe bestand darin, die Ziele über die Delegationsinteraktion schnellstmöglich auszuwählen und einem UAV zur Aufklärung zuzuweisen. Sobald ein UAV ein Foto des Objekts erstellt hat, flog es zum nächsten Ziel. Dazu konnten jedem UAV mehrere Ziele nacheinander zugewiesen werden und der Missionsplan gab dem Piloten eine Übersicht über die Anzahl und Reihenfolge der geplanten Aufklärungsaufträge. Außerdem war dort die Zeit ersichtlich, die für den Weg zum Ziel sowie das Erstellen der Aufnahme benötigt wird.

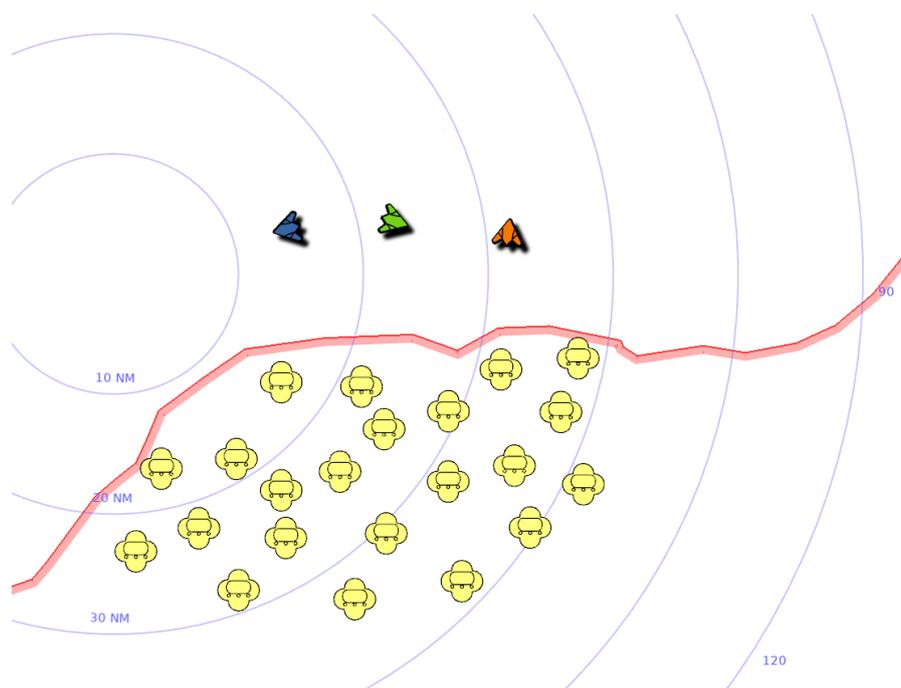


Abbildung 4-16 - Platzierung der UAVs und der auftauchenden Ziele zur Aufklärung im Missionsszenario 2

Den Probanden wurde der Auftrag gegeben, ein auftauchendes Ziel in einer möglichst kurzen Zeit zu klassifizieren. Die Klassifizierung eines Zieles erfolgte dabei durch die Probanden, die sich das erstellte Bild jedes Zieles ansehen und das Objekt anschließend als „hot“ für feindlich oder „cold“ für ungefährlich identifizieren mussten (siehe Anhang E.1). Der Versuch dauerte etwa 15 Minuten.

### 4.3.1.2 Messgrößen & Datenverarbeitung

Die aufgenommenen Messdaten entsprechen denen der Tabelle 4-3 aus dem vorhergehenden Versuch und wurden durch die SB-Bewertung, Leistungsmaße und Fragebögen ergänzt, welche im Folgenden beschrieben werden.

### SB-Modelle

Voraussetzung für die SB-Bewertung ist die Wissensbasis, die relevante Informationen im Sinne der Aufgabenstellung definiert. Die Relevanz von Informationen ist hier vorgegeben durch die Zielsetzung der Versuchsaufgabe. Abbildung 4-15 zeigt die Modelle für beide Versuche.

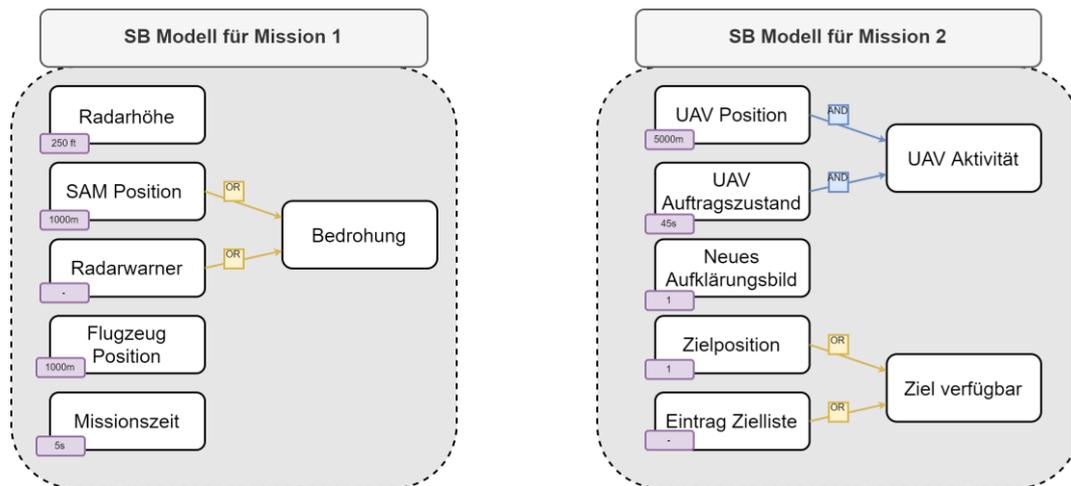


Abbildung 4-17 - SB-Modelle für Mission 1 (links) und Mission 2 (rechts). Die Normierungskonstante ist bei jedem SB-Knoten unten links annotiert (lila).

Das **SB-MODELL** des ersten Versuchs hat eine triviale Zuweisung von relevanten Informationselementen zu der Versuchsaufgabe. Die Radarhöhe wird im oberen rechten Rand des HUDs angezeigt (siehe Appendix Abbildung 1). Eine Bedrohung kann entweder über die taktische Karte oder über den Indikator des Radarwarners wahrgenommen werden, welcher sich unten rechts im HUD befindet. Der Indikator der Missionszeit ist in der Statusleiste des MHDDs zu sehen und die Flugzeugposition wird in der taktischen Karte dargestellt.

Das **SB-MODELL** des zweiten Versuchs beinhaltet zunächst die **SB-KNOTEN** „UAV Position“ und „UAV Auftragszustand“. Erstere wird auf der taktischen Karte dargestellt, auf der das Symbol des UAVs die assoziierte **OBSERVATION** erstellt. Der Auftragszustand ist im Missionsplan dargestellt und wird mit der aktuellen Missionszeit als **INHALT** befüllt. Die Wahl der Missionszeit als Inhalt modelliert somit, dass der Bediener nach 45 Sekunden ohne Fixation auf dem Missionsplan kein gutes SB vom Auftragszustand der UAV hat. Die Verbindung dieser beiden **SB-KNOTEN** mit der **AND-RELATION** modelliert, dass ein Bediener, der sowohl die UAV-Position als auch den Auftragszustand kennt, ein gutes SB der UAV-Aktivität hat. Neben den UAV-Informationen sind noch die Knoten „Neues Aufklärungsbild“, „Zielposition“ sowie „Eintrag Zielliste“ modelliert, welche jeweils den Informationenanzeigen auf der taktischen Karte beziehungsweise dem Eintrag in der Zielliste im MHDD entsprechen.

### Leistungsmaße

Im ersten Versuch dienten die Meldezeit bei einer Bedrohung sowie die Fehlerrate bei der Tiefflugaufgabe als Leistungsmaße. Die Meldezeit wurde definiert als die Zeit zwischen dem Auftauchen einer Bedrohung und dem Beginn der Meldung über Funk. Für die Fehlerrate im Tiefflug wurde die Zeit gemessen, die ein Versuchsteilnehmer über den vorgegeben 500 Fuß Radarhöhe flog und mit der Gesamtzeit normiert.

In der zweiten Mission wurden zwei Zeiten in der Durchführung der Delegations- und Klassifikationsaufgabe gemessen, welche im Zeitstrahl in Abbildung 4-18 abgebildet sind.

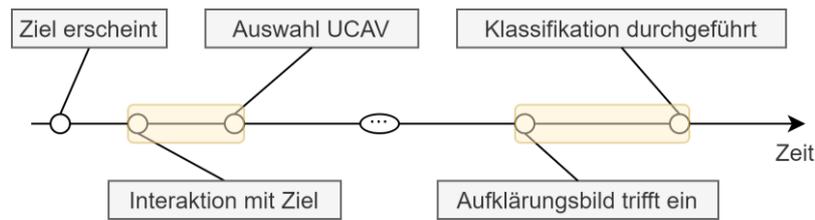


Abbildung 4-18 - Zeitstrahl der Aufgabe in Mission 2 mit gelb markierten Delegations- und Klassifikationszeit.

Die dem Versuch zugrundeliegende Annahme ist, dass ein Pilot mit gutem SB in der Lage ist, schneller das beste UAV auszuwählen und durch Antizipation eines Aufklärungsbildes auch rascher eine Klassifikation durchführt. Aus diesem Grund wurde die Delegationszeit

$$T_{Deleg}^{Ziel} = T_{Auswahl}^{UCAV} - T_{Auswahl}^{Ziel} \quad (8)$$

als Leistungsmaß definiert, die den Zeitraum zwischen Interaktion und Auswahl eines UAVs beschreibt. Analog dazu wurde die Klassifikationszeit

$$T_{Class}^{Ziel} = T_{Klassifiziert}^{Ziel} - T_{Bild\ eingetroffen}^{Ziel} \quad (9)$$

gemessen, welche die Zeit zwischen eintreffenden Aufklärungsbild und der durchgeführten Klassifikation beschreibt.

### Situation Awareness Rating Technique

Zur Messung des subjektiven Situationsbewusstseins wurde die SART nach der Beschreibung von Salmon et al. (2009) herangezogen. Daraus wurde ein Fragebogen übernommen und übersetzt. In diesem Fragebogen sind neun Fragen in den Kategorien benötigter und vorhandener Aufmerksamkeit sowie Verständnis. Jede Frage muss der Teilnehmer auf einer Skala zwischen eins und sieben beantworten. Der SB-Wert ergibt sich aus der Subtraktion der Antworten auf die Fragen bezüglich benötigter Aufmerksamkeit (Frage 1-3) von der Summe der Bewertungen in den Kategorien der vorhandenen Aufmerksamkeit (Frage 4-7) und dem Verständnis (Frage 8-9). In der Auswertung wird das relative Ergebnis verwendet, was der Normierung des Ergebnisses auf eine Zahl zwischen 0 und 1 entspricht. Der Fragebogen ist im Abschnitt E.3.1 im Anhang dargestellt.

### Datenverarbeitung

Während der Durchführung des Versuchs wurden die Messgrößen in eine Textdatei geschrieben und nach dem Versuch mit Python und den Bibliotheken *NumPy* und *Pandas* weiterverarbeitet. Die Diagramme wurden mit *Matplotlib* und *Seaborn* erstellt.<sup>12</sup> Zum Nachweis einer Normalverteilung in den Daten wurde der Shapiro-Wilk-Test angewendet (Shapiro & Wilk, 1965). Wenn der Test auf eine Normalverteilung der Daten hindeutet, wurde zur Angabe von Korrelation der Pearson Koeffizient verwendet. Wenn keine Normalverteilung angenommen werden kann, wurde der Korrelationskoeffizient nach Spearman berechnet (Spearman, 1961). Das Signifikanzniveau wurde auf  $p < 0.05$  festgesetzt. Zur Durchführung aller statistischen Tests wurde die Softwarebibliothek *SciPy* verwendet.

<sup>12</sup> Website: Matplotlib (<https://matplotlib.org/>), Seaborn (<https://seaborn.pydata.org/>),

### 4.3.1.3 Teilnehmer

Die Teilnehmer des Versuchs bestanden aus zwei Gruppen. Die erste Gruppe waren Kampfflugzeugpiloten der deutschen Luftwaffe. Tabelle 4-7 beinhaltet Details über die Teilnehmer des Versuchs.

ID	Alter	Typ (Flugstunden)	Flugstunden
P1	35	Tornado (700)	1000
P2	32	Tornado (920)	1200
P3	39	Tornado (1400)	1500
P4	35	WSO Tornado (1000)	1000
P5	51	F-4F (500) MiG-29 (1000 ) Eurofighter (2000)	3500
P6	38	F-4F (500) Eurofighter (200)	1160
P7	30	Eurofighter (220)	480
P8	30	Eurofighter (500)	1000

Tabelle 4-7 - Teilnehmergruppe der Kampfflugzeugpiloten mit dem Kürzel „P“

Die zweite Gruppe bestand aus neun Anwärtern des fliegerischen Dienstes der deutschen Luftwaffe, die sich zu diesem Zeitpunkt im Studium an der UniBw München befunden haben (mittleres Alter = 22,6 Jahre, Standardabweichung = 1,4, alle männlich). Diese Probanden werden im Folgenden mit „S“ und einer laufenden Nummer benannt.

### Genauigkeit der Blickbewegungsmessung

Eine Messung der Genauigkeit der Blickbewegungsmessung wurde nach der Kalibrierung des Blickbewegungsmesssystems durchgeführt. Hierfür wurden neun Punkte pro Bildschirm präsentiert und pro Punkt 100 Datenpunkte aufgenommen. Aus diesen Datenpunkten wurde dann die mittlere euklidische Distanz berechnet sowie die Standardabweichung der Stichprobe. Die Ergebnisse sind in Abbildung 4-19 für die zwei relevanten Bildschirme HUD und MHDD dargestellt.

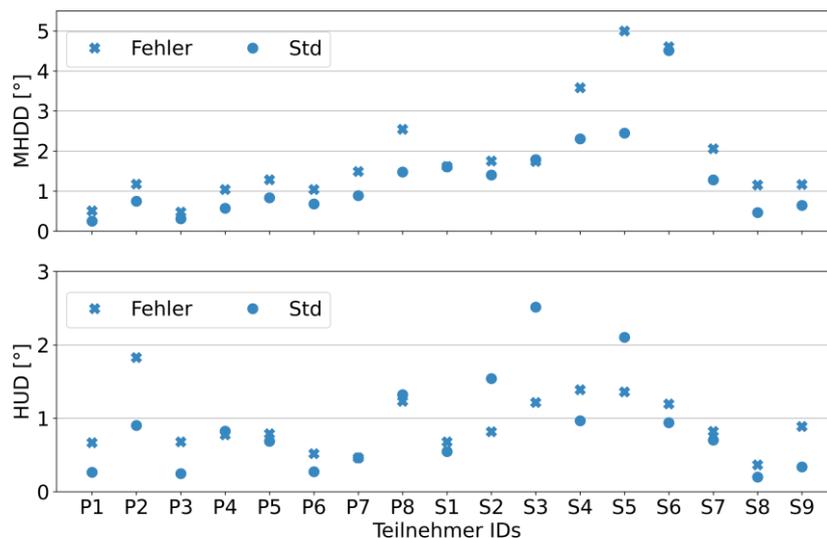


Abbildung 4-19 - Individuelle Genauigkeit der Blickbewegungsmessung quantifiziert durch den durchschnittlichen Fehler und Standardabweichung

Der Median der Winkelfehler liegt bei  $\tilde{x}_{MHDD} = 1.39^\circ$  sowie  $\tilde{x}_{HUD} = 0.8^\circ$ . Dass die Genauigkeit auf dem HUD höher ist, stimmt auch mit Erfahrungswerten überein. Grund hierfür ist, dass die Augen des Bedieners beim Blick auf die zentrale Außensicht so positioniert und geöffnet sind, dass die seitlich und unter dem Cockpitdisplays platzierten Kameras des Blickbewegungsmesssystems ein optimales Bild der Augen aufnehmen können. Es zeigen sich aber auch starke individuelle Unterschiede in der Messgenauigkeit, wie der Vergleich von P1 mit S5 in Abbildung 4-19 verdeutlicht.

#### 4.3.1.4 Ablauf

Die beiden Teilnehmergruppen hatten unterschiedliche Versuchsabläufe. Die Teilnehmergruppe der Kampfflugzeugpiloten machte den Versuch im Rahmen einer Experimentalwoche und wurde vor der Durchführung der Versuche zwei Tage lang trainiert. Dieses Training war nicht explizit auf diesen Versuch abgestimmt, die Piloten waren jedoch sehr vertraut mit der Bedienung des Simulators und dem Führen des Flugmodells. Diese Teilnehmergruppe erhielt kein spezifisches Training, sondern bekamen nur eine Erklärung in Form einer Präsentation über die anstehenden Aufgaben vor jeder Mission. Bevor die Teilnehmer dann den Versuch durchführten, wurde das Blickbewegungssystem kalibriert.

Die zweite Teilnehmergruppe der Studenten hatte keine Erfahrung mit dem Simulator. Diese Gruppe bekam ein dreistufiges Training, welches aus einer Simulatoreinführung und einem spezifischen Training für die Versuchsaufgabe bestand. Im spezifischen Teil wurde zunächst die Route des Versuchs ohne Flugabwehrstellung geflogen, anschließend eine äquivalente Route mit unbekanntem Flugabwehrstellungen.

#### 4.3.2 ERGEBNISSE

Im Folgenden werden die Ergebnisse getrennt nach Versuchen dargestellt, da sich die Auswertung auf die aufgabenspezifischen Informationen beziehen. Wie im vorherigen Abschnitt dargestellt, wurden zwei unterschiedliche SB-Modelle in den Versuchen verwendet.

##### 4.3.2.1 Mission 1

Im Folgenden werden die Ergebnisse der Tiefflug- und Meldeleistung sowie des SART-Fragebogens mit den im SB-Modell gemessenen Abweichungen in den für die Leistung relevanten Informationen verglichen. Anschließend werden Fehler im Zusammenhang mit der SB-Bewertung analysiert.

#### Leistung im Tiefflug

Zunächst soll der Zusammenhang zwischen der Fehlerrate eines Versuchsteilnehmers und der gemessenen Abweichung im Knoten „Radarhöhe“ des SB-Modells betrachtet werden. Hierfür zeigt die Abbildung 4-20 links die mittlere SB-Abweichung aufgetragen über die Fehlerzeit. Unter den Piloten gab es einen Teilnehmer, der eine wesentlich höhere relative Fehlerzeit als alle anderen Teilnehmer hatte. Um normalverteilte Daten annehmen zu können, wurde dieser als Ausreißer aus der weiteren Berechnung entfernt.

Nach Ausschluss dieses Datenpunkts wurde für die relative Fehlerzeit und die durchschnittlich gemessene Abweichung in der Radarhöhe ein signifikanter Korrelationskoeffizient von  $r=0.59$  ( $p<0.02$ ) berechnet. Eine höhere Fehlerzeit in der Tiefflugaufgabe ist also auch mit einer höheren Abweichung im **SB-KNOTEN** „Radarhöhe“ assoziiert. Der Ausreißer passt gut zum statistischen Trend der Ergebnisse der anderen Teilnehmer obwohl nicht der Teil der Berechnung ist. Es ist jedoch anzumerken, dass die Datenpunkte mit höheren Fehlerraten (P4, P5 & P6) den statistischen Trend stark beeinflussen und der Shapiro-Wilk-Test bei der Verteilung der Fehlerrate mittels an der Grenze zur Ablehnung einer

Normalverteilung liegt ( $p = 0.052$ ). Wenn die drei benannten Probanden mit hohen Fehlerraten ausgeschlossen werden, verschwindet auch die positive Korrelation.

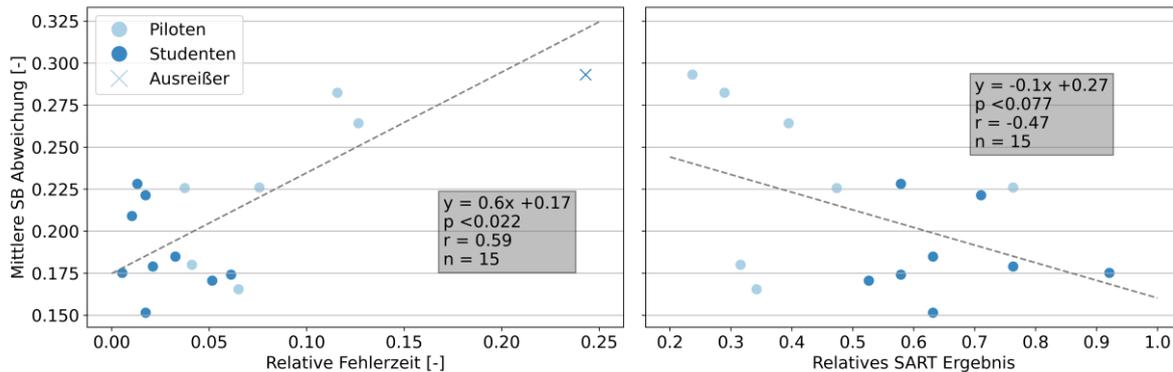


Abbildung 4-20 - Zusammenhang zwischen mittlerer Abweichung im Knoten Radarhöhe und der relativen Fehlerzeit sowie dem relativen SART-Ergebnis.

Die Daten legen einen Zusammenhang zwischen mittlerer Leistung und mittlerer SB-Abweichung nahe, sind jedoch nicht direkt nutzbar zur Verarbeitung in einem Assistenzsystem, da sie keine Informationen über konkrete Fehlersituationen beinhalten. Vergleichbar mit der Auswertung in Abschnitt 4.2.2 ist es hilfreich, die Vorhersage von Fehlern als Klassifikationsproblem zu formulieren. Im Folgenden wird die gemessene SB-Abweichung verwendet, um vorherzusagen, ob ein Fehler begangen wird und dies mit dem tatsächlichen Fehler abzugleichen.

Hierfür wurden zu jedem Zeitschritt der Messung die folgenden Klassifikationen vorgenommen: Lag die Radarhöhe über 500 Fuß, wurde dieser Zeitschritt als „Fehler“ klassifiziert, sonst als „Kein Fehler“. Lag die Abweichung über dem Grenzwert von  $dev_{RadAlt} = 1 - e^{-1} \approx 0.63$ , wurde der Zeitschritt als „Hohe Abweichung“ klassifiziert, sonst als „Niedrige Abweichung“. Abgeleitet aus dem dargestellten Zusammenhang zwischen SB-Abweichung und Fehler wird im Folgenden eine hohe Abweichung als Prädiktor für einen Fehler verwendet. Somit wird ein Zeitschritt als richtig-positiv klassifiziert, wenn er eine „Hohe Abweichung“ und einen „Fehler“ vorliegen hat. Die mittleren Ergebnisse aller Klassifikationen sind in der Tabelle 4-8 dargestellt. Auffällig ist zunächst, dass die Teilnehmer mit einer Ausnahme von P4 sehr niedrige Fehlerraten haben, weshalb die mittlere Positivrate bei  $p_{Fehler} = 0.114$  liegt („Positiv“ = Fehler). Von diesen wenigen Fehlern wurde bei 22.1% eine hohe Abweichung gemessen. In 77.9% der Fehlerdatenpunkte wurde keine hohe Abweichung gemessen. Zu diesen Klassifikationsfehlern kann es kommen, wenn ein Pilot die fehlerhafte Höhe bemerkt, aber noch nicht korrigiert hat. In den Datenpunkten, in denen kein Fehler vorlag, war die SB-Abweichung in 96.6% der Fälle niedrig, während bei 3.4% dieser Datenpunkte eine hohe Abweichung gemessen wurde, obwohl kein Fehler vorlag. Diese Art der Fehlklassifikation passiert, wenn ein Versuchsteilnehmer unter 500 Knoten fliegt, ohne die genaue Höhe zu kennen. Dies ist möglich, da der Abweichungsgrenzwert von über  $dev_{RadAlt} = 0.63$  einer abweichenden Radarhöhe von 250 Fuß entspricht und die Wahl des Normierungsparameters  $C_n$  von der Dynamik der Situation aber auch von dem Ziel abhängt, das ein Bediener verfolgt. Es ist ein Unterschied, ob ein Pilot innerhalb eines Höhenbands von 300 Fuß oder 1000 Fuß fliegen muss.

	Fehler	Kein Fehler
<b>Hohe Abweichung</b>	0.221 ( $\pm 0.147$ )	0.034 ( $\pm 0.017$ )
<b>Niedrige Abweichung</b>	0.779 ( $\pm 0.147$ )	0.966 ( $\pm 0.017$ )
<b>Prävalenz</b>	0.114 ( $\pm 0.120$ )	0.886 ( $\pm 0.120$ )

Tabelle 4-8 - Mittlere Klassifikationsrate angegeben in Durchschnitt (Standardabweichung). Durchschnittliche Anzahl der Datenpunkte  $\bar{n} = 15612$ .

Die individuellen Klassifikationsraten sind in Abbildung 4-21 dargestellt. Dabei wurde auf falsch-positive und falsch-negative Datenpunkte verzichtet, da diese stets invers zur richtig negativen beziehungsweise richtig positiven Rate liegen. Zusätzlich ist dort die Positivrate eingetragen, also die Anzahl der Datenpunkte, bei denen ein Fehler in der Radarhöhe vorlag.

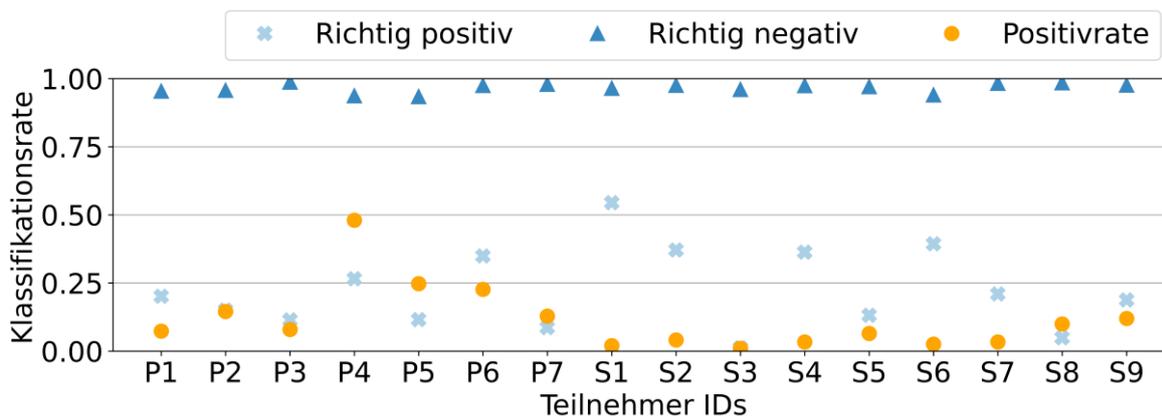


Abbildung 4-21 - Klassifikationsleistung und Positivrate über den individuellen Teilnehmern (P=Pilot, S=Student)

Die individuellen Daten zeigen, dass es über alle Probanden hinweg kaum Situationen gab, bei denen eine hohe SB-Abweichung gemessen wurde und kein Fehler vorlag. Bei der Klassifikation von richtig positiven Klassifikationen (Fehler & hohe SB-Abweichung) gibt es durchaus individuelle Unterschiede. Die Gründe für diese unterschiedlichen Raten könnten darin liegen, dass die Piloten individuell schnell auf einen Fehler im Tiefflug reagierten und auch das Flugzeug mit unterschiedlichem Geschick zurück unter die 500 Fuß steuern konnten. Einen großen Anteil an den falsch positiven Fällen haben die Zeiträume, in denen der Proband einen Fehler korrigiert, nachdem er ihn festgestellt hat.

## SART

Die Abbildung 4-20 zeigt rechts den Zusammenhang zwischen der subjektiven Einschätzung des Situationsbewusstseins und der SB-Abweichung. Hier zeigt sich ein nicht signifikanter negativer Pearson-Korrelationskoeffizient von  $r = -0.42$  von durchschnittlicher SB-Abweichung in der Tiefflugaufgabe mit dem SART-Bewertungsergebnis. Dies bedeutet, dass die Versuchsteilnehmer selbst merken, dass es einen Zusammenhang zwischen der subjektiven Wahrnehmung des eigenen SB und der gemessenen SB-Abweichung gibt. Möglicherweise stellten die Probanden während der Versuchsdurchführung fest, dass sie einer Information zu lange keine Beachtung geschenkt haben und Fehler entstanden sind. Vor allem in der Tiefflugaufgabe war es möglich, anhand der Radarhöhenanzeige unmittelbar einen Fehler festzustellen. Bei diesem Ergebnis gilt es aber zu beachten, dass die Meldeaufgabe mit hoher Wahrscheinlichkeit einen Einfluss auf das Ergebnis des Fragebogens hatte und in der Korrelationsberechnung bisher nicht beachtet wurde.

## Meldeleistung

Im Folgenden wird sich der Meldeleistung gewidmet, also inwiefern die SB-Abschätzung auch vorhersagen konnte, dass ein Pilot eine Bedrohung wahrgenommen hat. Die Abbildung 4-22 zeigt die Verteilung der Zeitdifferenz zwischen dem Meldezeitpunkt des Versuchsteilnehmer und dem Moment, in dem die SB-Abschätzung die Wahrnehmung der spezifischen Bedrohung registriert hat. Dies wurde über den **SB-KNOTEN** „Bedrohung“ umgesetzt, welcher für jede individuelle Bedrohung erstellt wird und die **IDENTIFIKATION** der SAM trägt.

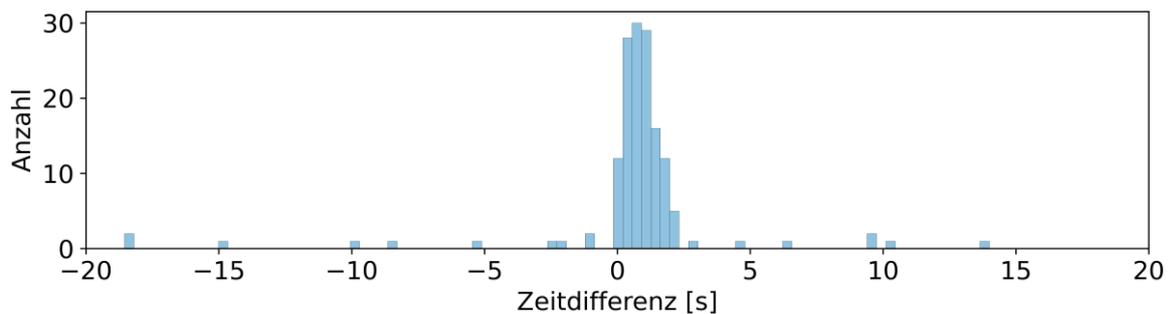


Abbildung 4-22 - Zeitdifferenz zwischen Meldezeitpunkt und dem Zeitpunkt der ersten  $dev_{\text{Bedrohung}} = 0$  im Knoten „Bedrohung“ der zugehörigen SAM-Site-Identifikation (n=157)

Eine Meldung wurde als richtig klassifiziert bezeichnet, wenn der Pilot innerhalb von drei Sekunden nach der ersten im Knoten registrierten Wahrnehmung der Bedrohung eine Meldung abgesetzt hat. Die Abweichungsmessung hat in 82% ( $n = 130$ ) aller 157 Meldungen die Wahrnehmung der Bedrohung richtig klassifiziert – in drei Fällen der möglichen 160 Datenpunkte wurde eine falsche Meldung abgesetzt, weshalb diese Datenpunkte ausgeschlossen wurden. In 21 Fällen wurde eine Meldung abgesetzt, ohne dass im Messsystem die Wahrnehmung der entsprechenden Bedrohung gemessen wurde. In sechs Fällen wurde die Meldung erst nach drei Sekunden nach der Wahrnehmung abgesetzt. Dies kann an einem Messfehler liegen oder an der kurzzeitigen Priorisierung des Tiefflugs über diese Aufgabe. Mit diesem Ergebnis kann auch demonstriert werden, dass die Abschätzung von SB über identifizierbare Objekte nützlich sein kann. Hier wird die Unterscheidung von „Bedrohung“ nicht allgemein getroffen, sondern mit Bezug auf eine individuelle SAM-Site, was die **IDENTIFIKATION** im **SB-MODELL** möglich macht.

## Fehlleistung

Die bisher vorgestellten Zusammenhänge legen einen grundsätzlichen Zusammenhang zwischen der SB-Abschätzung und dem tatsächlichen SB der Versuchsteilnehmer nahe. Die Anwendung in einem Assistenzsystem bedingt aber, dass Lücken im SB des Bedieners in konkreten Fehlersituationen festgestellt werden können. Daher soll im Folgenden ein Fokus auf Situationen gelegt werden, in denen Fehler passiert sind und die Frage, ob die SB Abschätzung Indikatoren liefert, die diese Situationen identifizieren können.

Der erste Fehler im vorliegenden Versuch ist ein längeres Überfliegen der vorgegeben Radarhöhe von 500 Fuß. Hierfür wurden für jeden Versuchsteilnehmer die Situation ausgewählt, bei der am längsten am Stück über dem Grenzwert der Tiefflugaufgabe geflogen wurde. Für diesen Zeitraum wurde in Abbildung 4-23 und Abbildung 4-24 der Verlauf der Abweichung für den Knoten „Radarhöhe“ sowie die tatsächliche Radarhöhe aufgetragen. Der stufenweise Verlauf der Abweichungswerte ist hier der niedrigen Erstellungsrate der Systemobservationen geschuldet.

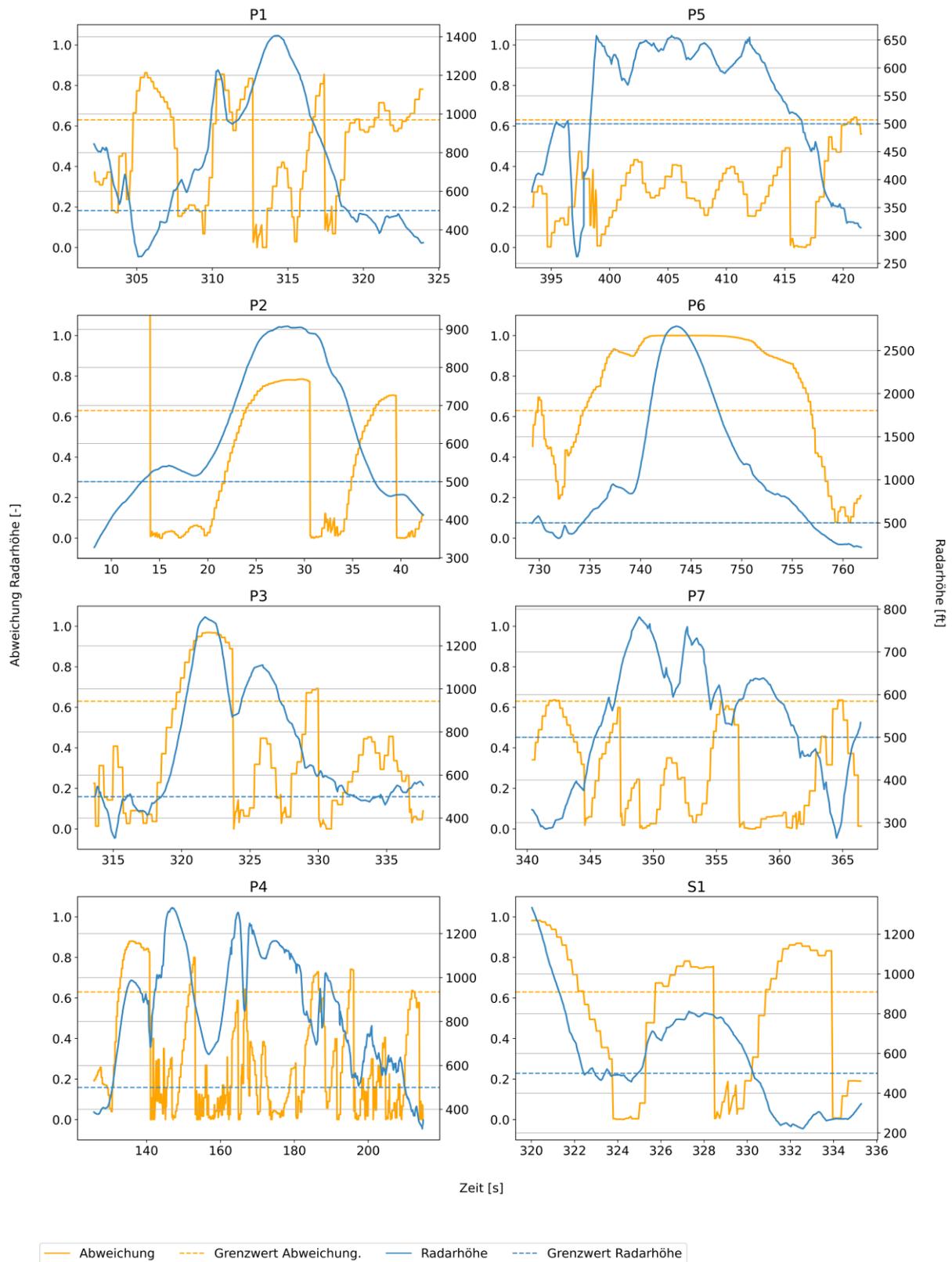


Abbildung 4-23 - Abweichung und Radarhöhe zum Zeitpunkt ausgewählter Fehlerfälle in Mission 1 für die Versuchspersonen P1-7 und S1. Bei P2 liegt ein Messfehler vor. Es gilt zu beachten, dass die individuelle y-Achsen in der Radarhöhe unterschiedlich sind, da die Fehlerhöhen unterschiedlich groß sind.

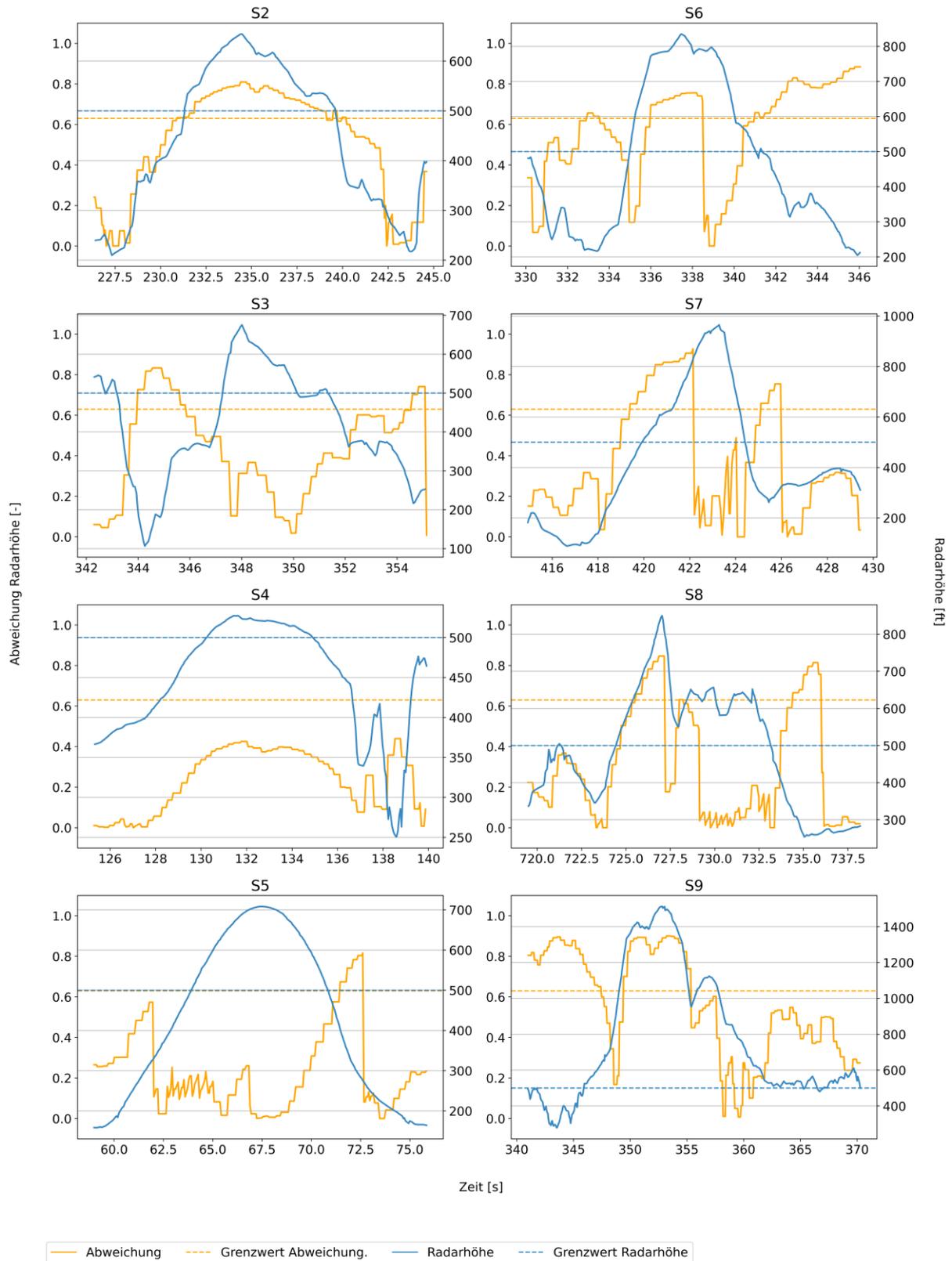


Abbildung 4-24 - Abweichung und Radarhöhe zum Zeitpunkt ausgewählter Fehlerfälle in Mission 1 für die Versuchspersonen S2-S9.

Zu Beginn von 11 der 16 Situationen<sup>13</sup> ähnelt der Verlauf der SB-Abweichung dem Verlauf der Radarhöhe, das heißt die SB-Abweichung steigt in dem Moment an, in dem die Probanden die Höhenbegrenzung übersteigen. Im nachfolgenden Verlauf ist es meist so, dass der Pilot die Radarhöhenanzeige fixiert, die Abweichung auf  $x_{dev}^{Radarhöhe} = 0$  springt und eine Korrektur der Flughöhe eingeleitet wird. Interessanterweise steigt bei einigen die Abweichung im Verlauf der Korrektur an<sup>14</sup>, da während der Korrektur nicht direkt die Flughöhe überprüft wird und somit die SB-Abweichung in SB-Knoten „Radarhöhe“ steigt, obwohl sich die tatsächliche Flughöhe in die richtige Richtung entwickelt.

Aus diesen Ergebnissen kann geschlossen werden, dass die situative Auswertung der Abweichung nicht immer ausreicht, um eindeutig Fehler in der Tiefflugaufgabe zu erkennen, eine ansteigende Abweichung aber oft in Kombination mit einem Fehler auftritt. Dies entspricht auch den Ergebnissen in der mittleren Klassifikationsleistung (siehe Tabelle 4-8), da in der situativen Betrachtung die gleichen Fehlklassifikationen gemacht werden. Ein bereits genannter Grund für diese fehlerhaften Klassifikationen ist, dass die Probanden ihre Tiefflugaufgabe auch gut erfüllen können, ohne den exakten Höhenwert abzulesen. Einerseits liegt dies am peripheren Sehen und der räumlichen Orientierung im Terrain, an dem die Probanden die Änderung ihrer Flughöhe abschätzen konnten. Darüber hinaus ist die kontinuierliche Überwachung der Flughöhe nicht notwendig, weil die Probanden die Zeitdauer von Übergängen der Flughöhe abschätzen können und ganz im Sinne von Wickens *Expectancy* abschätzen, zu welchem Moment die Höheninformation wieder relevant ist.

Der zweite Fehlertyp ist ein verspätetes Melden einer feindlichen Luftabwehrstellung. Basierend auf der Verteilung der Meldezeiten (siehe Abbildung 4-25) wurde nach Sichtung der Daten in diesem Versuch eine Dauer von  $T = T_{Meldung} - T_{Popup} > 6s$  als Fehler festgelegt, was für 11 Meldungen zutrifft.

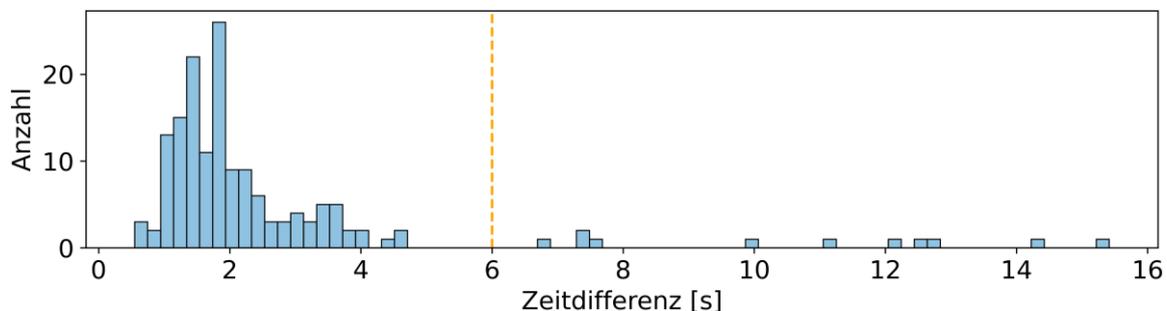


Abbildung 4-25 - Zeitdifferenz zwischen Auftauchen einer Bedrohung der Meldung. In orange markiert ist der Grenzwert, ab dem eine Zeitdifferenz als Fehler klassifiziert wird.

Bei 5 dieser 11 verspäteten Meldungen wurde im Zeitraum der ersten sechs Sekunden nach Auftauchen der Bedrohung im entsprechenden „Bedrohung“-Knoten eine Abweichung von  $x_{dev}^{Bedrohung} = -1$  gemessen. Bei 4 dieser 5 richtig positiven Klassifikationen wurde dann kurz vor der späten Meldung eine Wahrnehmung der Information gemessen.

Abbildung 4-26 zeigt die Messung der Abweichung in den vier relevanten **SB-KNOTEN** für einen dieser Fehlerfälle von P1 bei der zweiten Bedrohung, im Folgenden als „Bedrohung#2“ bezeichnet. Die SAM-Site erscheint zum Zeitpunkt  $T_{Popup} = 163.8s$  und der Pilot braucht über 14 Sekunden zur Meldung.

<sup>13</sup> Zu sehen bei P1-3, P6, S1, S2, S4, S6-8, S9

<sup>14</sup> Zu sehen bei P1-5, P7, S1, S5-8

Dies kann in der SB-Bewertung nachvollzogen werden, da die SB-Abweichung im Knoten „Bedrohung#2“ erst zum Zeitpunkt  $T_{Rec} = 176.9s$  auf  $x_{dev}^{Bedrohung\#2}(T_{Rec}) = 0$  springt und somit eine Wahrnehmung gemessen wurde. Um die geforderten Informationen durchzugeben, fixiert der Pilot kurz drauf die Anzeige für die Missionszeit.

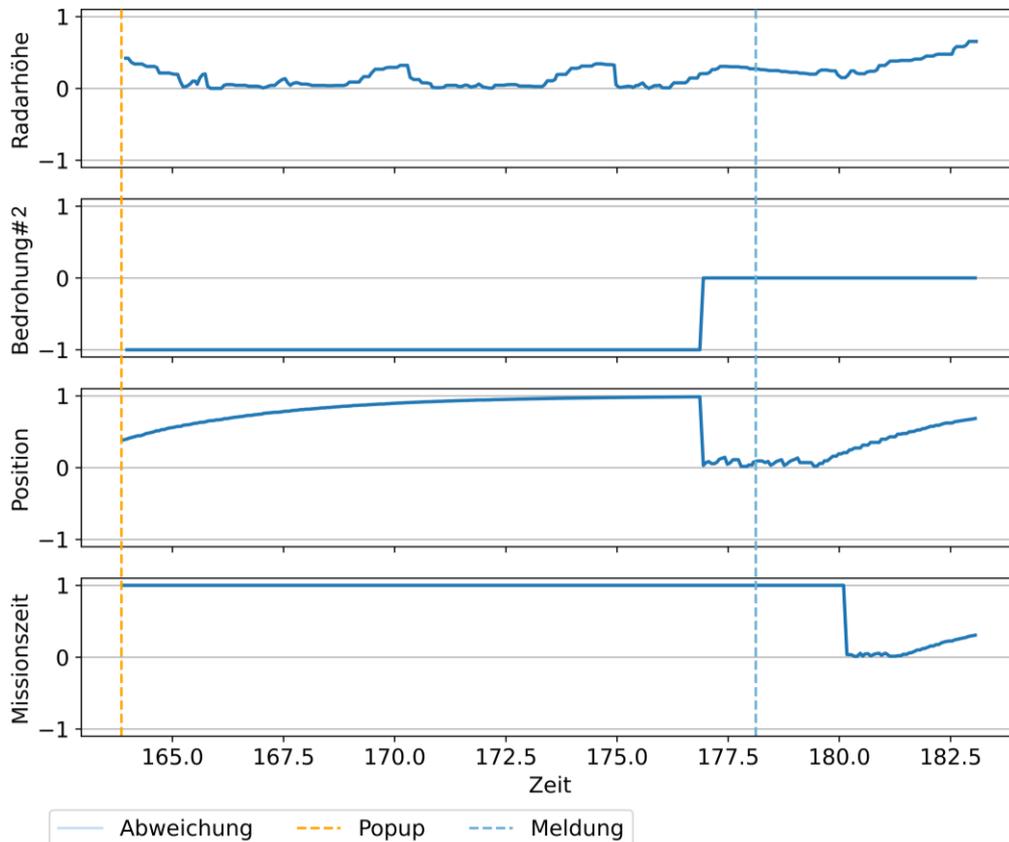


Abbildung 4-26 - Abweichungswerte von P1 in den Knoten „Missionszeit“, „Position“, „Bedrohung#2“ und „Radarhöhe“ für den Zeitraum ab Auftauchen der Bedrohung#2 (orange, gestrichelt) bis kurz nach Meldezeitpunkt (blau, gestrichelt).

Zum Vergleich zeigt Abbildung 4-27 die Abweichungswerte einer schnellen Meldung. Hier liegt der Zeitraum zwischen Auftauchen der Bedrohung  $T_{Popup} = 100.8s$  und der Meldung bei  $T_{Report}$  unter zwei Sekunden. An den Messwerten kann abgelesen werden, dass der Pilot zum Zeitpunkt  $T_{Rec} = 101.4s$  die Bedrohung bemerkt und dann zum Zeitpunkt  $T_{Report} = 102.6s$  eine Meldung absetzt. Hier ist ein ähnliches Muster zu erkennen, bei dem erst die Bedrohung wahrgenommen wird, dann die Position des Flugzeugs und zuletzt die Missionszeit, was sich zum Zeitpunkt  $T_{MT} = 106.2s$  in dem Sprung der Abweichung auf  $x_{dev}^{Missionszeit}(T_{MT}) = 0$  äußert.

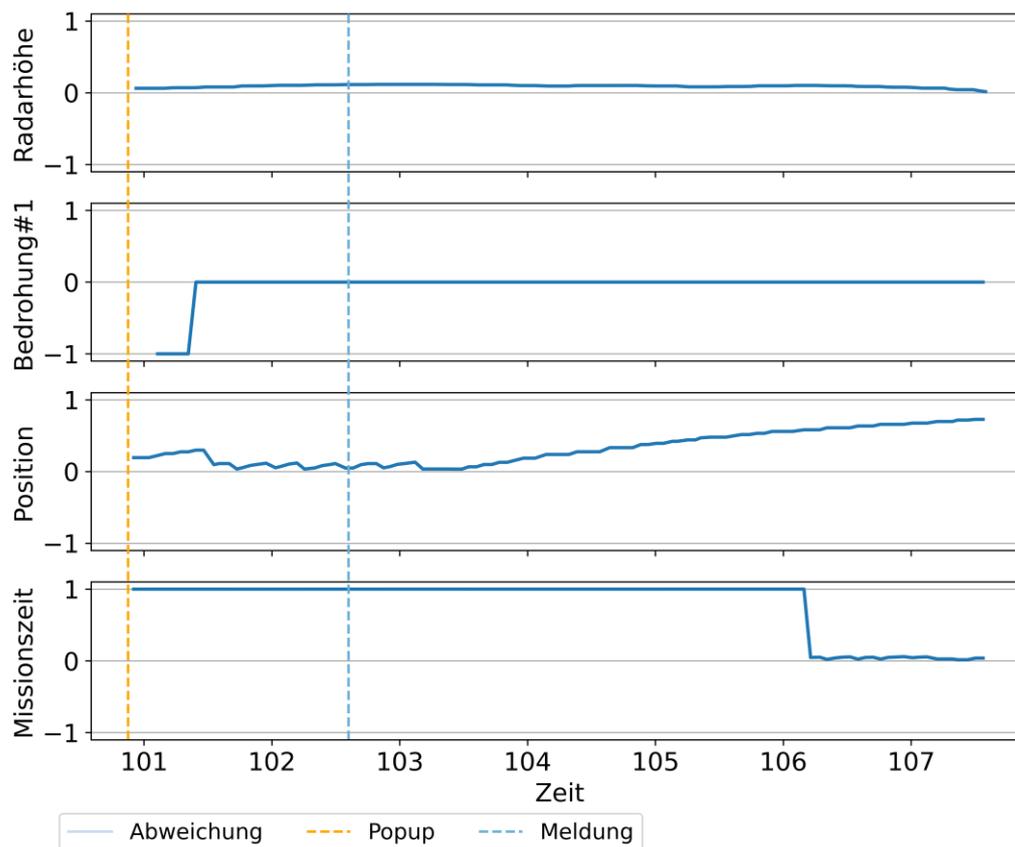


Abbildung 4-27 - Abweichungswerte von P1 in den Knoten „Missionszeit“, „Position“, „Bedrohung#1“ und „Radarhöhe“ für den Zeitraum ab Auftauchen der Bedrohung#1 (orange, gestrichelt) bis kurz nach Meldezeitpunkt (blau, gestrichelt)

Anhand dieser beiden Situationen kann auch demonstriert werden, warum das Einbeziehen der Systeminformationen in die SB-Abschätzung ein wichtiger Aspekt ist, um das Situationsbild eines Bedieners zu quantifizieren. Wie eben beschrieben verfolgen die Versuchsteilnehmer nach Bemerkung einer Bedrohung meist das Muster, erst die eigene Position auf der taktischen Karte und dann die Anzeige für die Missionszeit anzublicken. Da sich diese Informationen auf dem MHDD im Cockpit befinden, kann der Versuchsteilnehmer in diesem Zeitraum nicht seine Radarhöhe überwachen. Dieses Vernachlässigen der Radarhöhe kann unkritisch im Sinne der Tiefflugaufgabe sein, wenn das Terrain in diesem Zeitraum eben ist und sich nicht ändert. Es könnte aber auch kritisch sein, wenn das Terrain zu diesem Zeitpunkt stark variiert. Eine Abschätzung von SB rein auf Blickbewegungsdaten könnte zwischen diesen beiden Fällen nicht unterscheiden. Das hier vorgeschlagene Konzept misst dagegen unterschiedliche Abweichungswerte, was sich beim Vergleich der Radarhöhen-Abweichung in Abbildung 4-26 und Abbildung 4-27 zeigt. Die erste Situation in Abbildung 4-26 zeigt, wie die Radarhöhenabweichung stark steigt, während der Versuchsteilnehmer die Radarhöhe ignoriert, während dieser Abweichungswert in der zweiten Situation in Abbildung 4-27 stabil niedrig bleibt. Mit diesem Vergleich kann beispielhaft dargelegt werden, warum SB stets abhängig von der Situation ist und in die SB-Messung mit einfließen muss.

#### 4.3.2.2 Mission 2

Das Ziel der zweiten Mission war es, auftauchende Objekte möglichst schnell zu klassifizieren, also erst einen Aufklärungsauftrag an das beste UAV zu delegieren und dann die erstellten Bilder zu sichten. Die definierten Leistungsmaße bezogen sich auf die Zeit zwischen der Auswahl eines neuen Ziels und der Zuweisung eines UAVs (Delegationszeit) sowie dem Eintreffen eines neuen Bilds und dessen Klassifikation (Klassifikationszeit). Für die Auswertung wurde der **SB-KNOTEN** „UAV Aktivität“ zur Auswertung herangezogen, da er erstens zentral ist, um zu entscheiden welches UAV die beste Wahl für ein neues Ziel ist und zweitens damit auch das Eintreffen neuer Aufklärungsbilder antizipiert werden kann. Neben diesem Knoten wurde zur Fehleranalyse auch das Bemerkten von neuen Zielen der **SB-KNOTEN** „Neues Ziel verfügbar“ sowie „Neues Aufklärungsbild“ ausgewertet.

#### Geschwindigkeit von Delegation und Klassifikation

Im Versuch gab es drei UAVs und damit auch drei Abweichungswerte. Der Wert der im folgenden verwendeten SB-Abweichung ist der Durchschnitt dieser drei Abweichungen. Die Abbildung 4-28 zeigt den Zusammenhang zwischen der Abweichung im Knoten „UAV Aktivität“ und den mittleren Delegations- und Klassifikationszeiten pro Versuchsteilnehmer. Um für die Daten eine Normalverteilung annehmen zu können, wurden vier Versuchsteilnehmer bei der Delegationszeit ausgeschlossen, da diese durchschnittlich eine Delegationszeit von  $T_{Deleg}^{Ziel} > 8s$  hatten. Bei der Klassifikationszeit wurde ein Proband ausgeschlossen mit einer durchschnittlichen  $T_{Class}^{Ziel} > 40s$ . Nach Ausschluss dieser Daten kann laut dem Shapiro-Wilk-Test eine Normalverteilung angenommen werden. Bei den verbleibenden Datenpunkten korreliert die Delegationszeit positiv mit der Delegationszeit mit einem Koeffizienten von  $r = 0.59$ . Bei der Klassifikationszeit ist die positive Korrelation noch stärker ausgeprägt mit einem Koeffizienten von  $r = 0.77$  und ebenfalls statistisch signifikant. Die Ergebnisse sprechen dafür, dass alle Bediener, die einen Überblick über alle UAV-Positionen und Aufträge behalten haben, auch eine bessere Leistung in dieser Aufgabe vorweisen konnten.

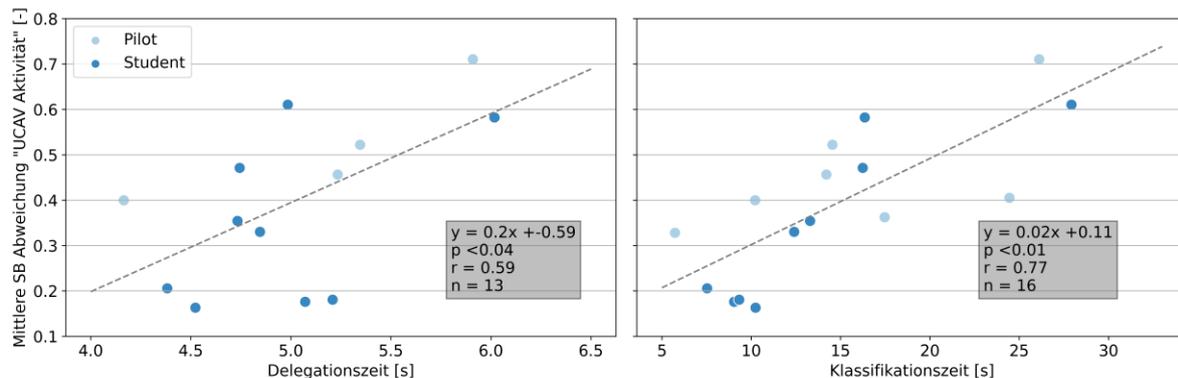


Abbildung 4-28 - Zusammenhang zwischen mittlerer Delegationszeit und mittlerer SB-Abweichung im Knoten „UAV-Aktivität“

#### SART

Ähnlich wie beim ersten Versuch ist in Abbildung 4-29 ein negativer Zusammenhang zwischen den Ergebnissen der SB-Selbstbewertung und der mittleren SB-Abweichung im Knoten „UAV Aktivität“ sichtbar. Der Korrelationskoeffizient liegt in diesem Fall bei  $r = -0.5$  und ist statistisch nicht signifikant. Diese Ergebnisse legen die Vermutung nahe, dass sich die Versuchsteilnehmer bewusst

sind, dass sie Informationen vernachlässigt haben oder dass sie von der Dynamik der Situation überrascht werden, was von der SB-Bewertung gemessen werden konnte.

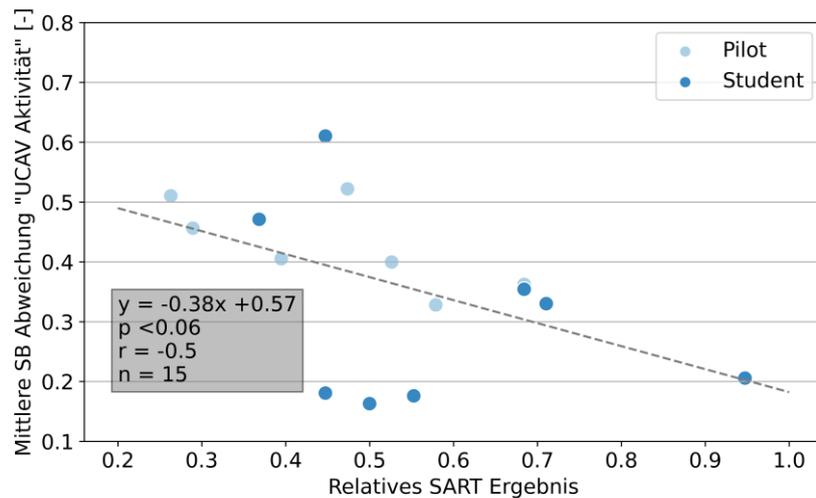


Abbildung 4-29 - Zusammenhang zwischen relativem SART-Ergebnis und mittlerer SB-Abweichung im Knoten „UCAV Aktivität“

## Fehlleistung

Die Ergebnisse des zweiten Versuchs deuten auf einen Zusammenhang zwischen mittleren Leistungsmaßen, SART und der SB-Abschätzung hin. Wie auch bei der Tiefflugaufgabe gilt hier, dass der bestätigte Zusammenhang nur schwer in einem Assistenzsystem zum Einsatz kommen kann. Hierfür sind die Fehler interessant, die außerhalb der durchschnittlichen Leistungen in einer Aufgabe liegen, also die statistischen Ausreißer (engl. *Outlier*). Im zweiten Versuch können Fehler beim Bemerkens neuer Ziele, bei der Auswahl der UAVs sowie bei der Bearbeitungsdauer von Bildern passieren. Für aufgetretene Fehlerfälle, die besonders gravierend waren, soll im Folgenden der Zusammenhang mit der SB-Abschätzung analysiert werden. Dabei werden drei verschiedene Fehler analysiert: Das Übersehen eines neuen Ziels, das Zuweisen einer Aufgabe an das falsche UAV und das Übersehen eines neuen Aufklärungsbilds.

### Übersehen eines Ziels

Ein Ziel wurde übersehen, wenn die Zeit zwischen Auftauchen und Interaktionsbeginn außerordentlich lang ist. Abbildung 4-30 zeigt die Verteilung der Zeitdifferenzen aller Piloten und den Grenzwert  $T_{Popup,Select}^{Ziel} = 20s$ , ab dem die Zeitdifferenz als Fehler gewertet wurde. Für alle Ziele, bei denen die Zeitdifferenz über dem festgelegten Grenzwert liegen, zeigt Abbildung 4-31 die Zeitdifferenzen zwischen der ersten gemessenen Wahrnehmung und der Auswahl des Ziels. Bei 36.7% ( $n=26$ ) dieser 71 Fehlerfälle konnte gemessen werden, dass die Aktion innerhalb von  $T=3s$  auf die erste gemessene Wahrnehmung folgte. Dabei muss aber angemerkt werden, dass 20 der 26 Messungen durch die direkte Interaktion mit dem entsprechenden Objekt auf der taktischen Karte gemessen wurde und nicht durch die Blickbewegungsmessung. Die zeigt, dass die Anforderung an die Blickbewegungsmessung auf der taktischen Karte besonders hoch ist und das Auswerten eines Radius von  $2^\circ$  um den gemessenen Fixationspunkt auch zu vielen falsch positiven Ergebnissen führt. Dies äußert sich hier vor allem in der hohen Zahl an Delegationen mit hoher Zeitdauer zwischen gemessener Wahrnehmung und Delegation. Ein zweiter Grund ist auch, dass die Probanden teilweise abwarteten bis ihr UAVs in der Missionsausführung etwas vorangeschritten waren, um neue Aufträge zuzuweisen. Diese Strategie war vorteilhaft im Sinne der Gesamtaufgabe, da zu einem späteren

Zeitpunkt meist neue Ziele bekannt waren. Für das hier angenommene Leistungsmaß war dieses Verhalten dagegen ein Störfaktor.

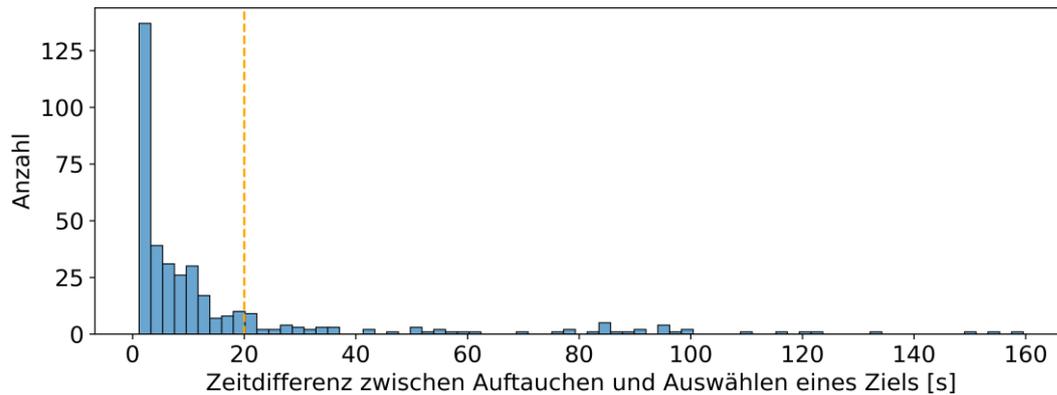


Abbildung 4-30 - Zeitdifferenz zwischen Auftauchen und Auswahl eines Ziels. In orange markiert ist der Grenzwert, ab dem eine Zeitdifferenz als Fehler klassifiziert wird ( $T > 20$ ).

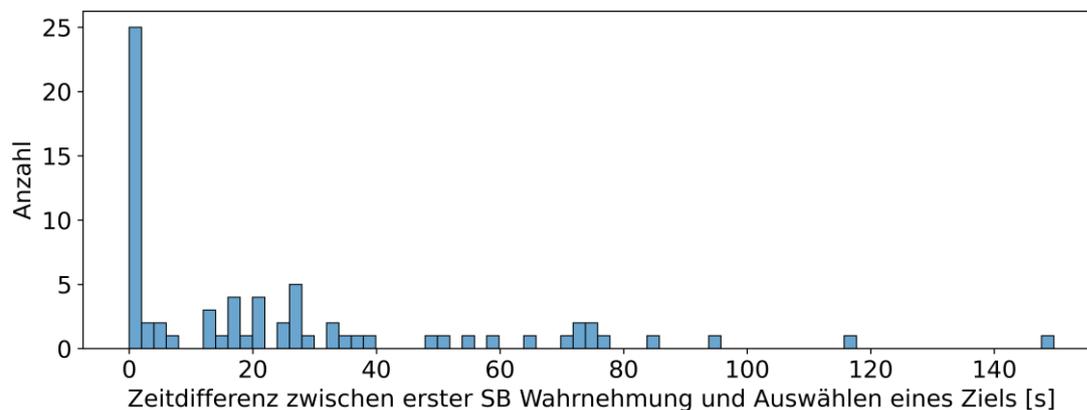


Abbildung 4-31 - Zeitdifferenz zwischen dem Zeitpunkt der gemessenen Wahrnehmung im SB-Modell und der Auswahl eines Ziels für alle Datenpunkte mit Fehler.

#### Falschwahl eines UAVs

Zentrale Aufgabe bei Auftauchen eines neuen Ziels war die Auswahl des korrekten UAVs im Sinne der Optimierung der Missionszeit. Von den insgesamt 374 getroffenen Entscheidungen waren 46 inkorrekt. Die mittleren Abweichungen für die korrekten Entscheidung betragen  $dev_{Pos,UCAVs}^{Korrekt} = 0.134$  und  $dev_{Activity,UCAVs}^{Korrekt} = 0.175$  und für inkorrekte Entscheidungen  $dev_{Pos,UCAVs}^{Inkorrekt} = 0.119$  und  $dev_{Activity,UCAVs}^{Inkorrekt} = 0.213$ . Die Unterschiede in der Abweichung sind hierbei so gering, dass sie sich nicht für eine Klassifikation eignen. In dieser Aufgabe reicht die SB-Bewertung nicht aus, um eine Fehlleistung präzisieren zu können. Beobachtung und Nachbesprechungen mit den Probanden offenbarten, dass die Probanden strategische Entscheidungen gegen eine optimale Verteilung eines neuen Ziels an die UAVs trafen, um UAVs möglichst gleichmäßig im Missionsgebiet zu verteilen. Diese Strategie konnte in der Gesamtmission tatsächlich zu einem besseren Zeitergebnis führen, war aber so in der Auswertung nicht vorgesehen.

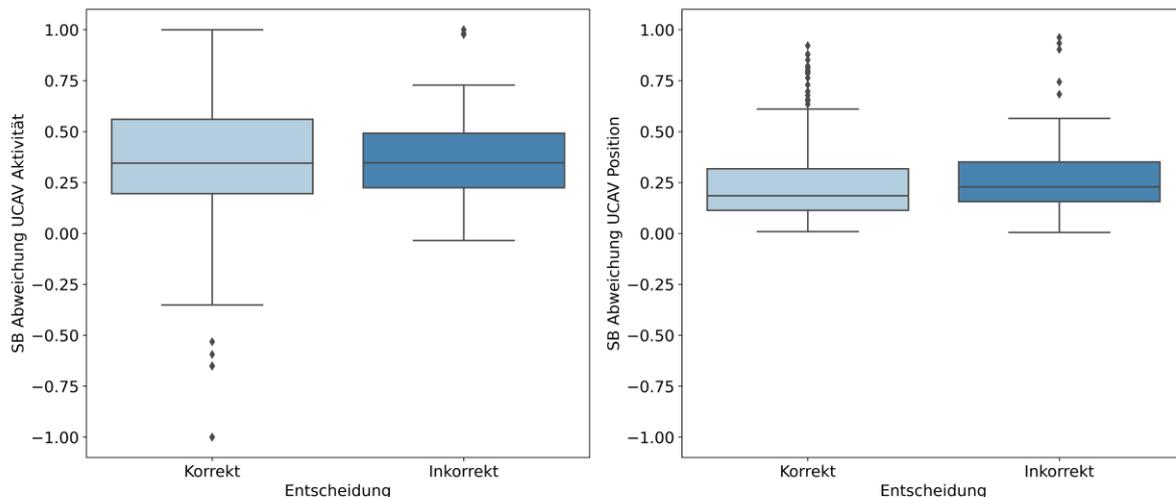


Abbildung 4-32 - Abweichungswerte in den Knoten „UCAV Aktivität“ und „UCAV Position“ für korrekte und inkorrekte Entscheidungen.

### Übersehen eines neuen Aufklärungsbilds

Ein Aufklärungsbild wurde übersehen, nachdem ein UAV einen Aufklärungsauftrag durchgeführt hat und die Klassifikation des Bilds länger als ein definierter Grenzwert dauerte. Die Abbildung 4-33 zeigt die Verteilung der Zeitdifferenzen aller Probanden zwischen Auftauchen eines Aufklärungsbild und dem Beginnen des Klassifikationsprozesses. Dort ist auch der Grenzwert  $T = 40s$  markiert, ab dem ein Klassifikationsvorgang als Fehler gewertet wurde.

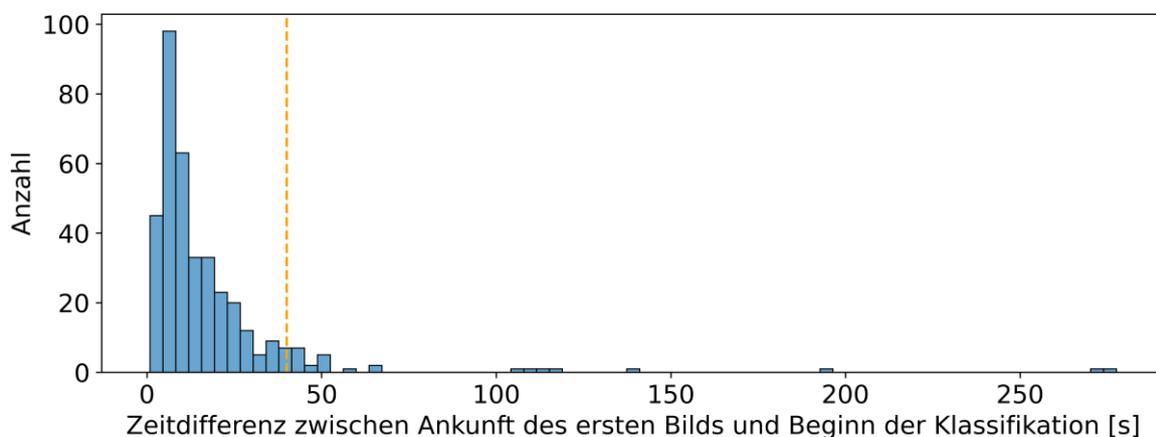


Abbildung 4-33 - Zeitdifferenz zwischen Auftauchen und Auswahl eines Aufklärungsbilds. In orange markiert ist der Grenzwert, ab dem eine Zeitdifferenz als Fehler klassifiziert wird ( $T > 40$ ).

Ein Blick auf die Zeitdifferenz zwischen der gemessenen Wahrnehmung durch die das Modell im **SB-KNOTEN** „Neues Aufklärungsbild“ und dem tatsächlichen Beginnen des Aufklärungsvorgangs in Abbildung 4-34 zeigt, dass ein großer Teil der 95 Zeitdifferenzen unter  $T = 3s$  liegen ( $n=43$ ). Das bedeutet, dass in der langen Zeitdauer vor Beginnen der Klassifikation eine hohe Abweichung in dem Knoten „Neues Aufklärungsbild“ mit der entsprechenden Identifikation vorlag. Wie bei der Delegation ist es auch hier so, dass ein Großteil ( $n=41$ ) dieser Fälle durch die Bedienung der mit der Klassifikation im Zusammenhang stehenden Bedienelemente gemessen wurde, weshalb der Zeitpunkt der Wahrnehmungsmessung mit dem der Interaktion zusammenfällt und damit bei  $T = 0s$  liegt.

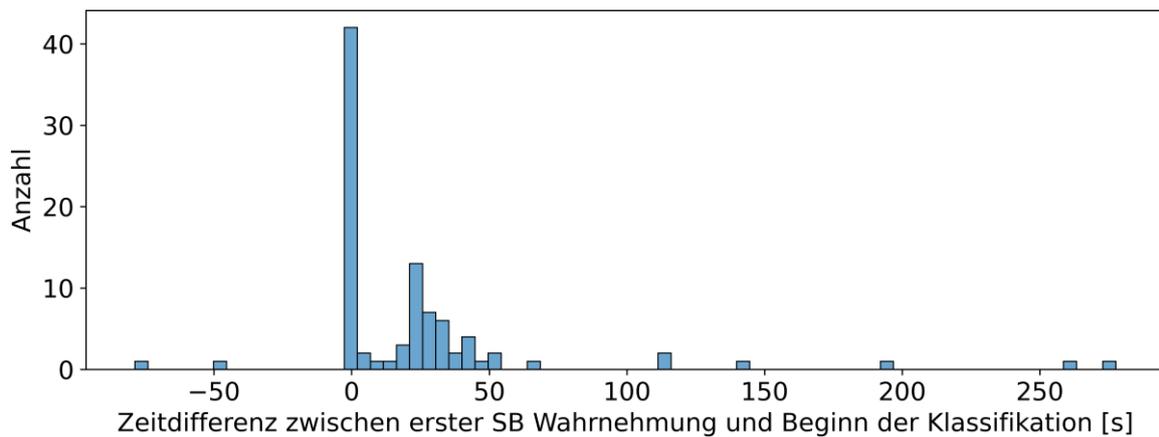


Abbildung 4-34 - Zeitdifferenz zwischen dem Zeitpunkt der gemessenen Wahrnehmung im SB-Modell und dem Beginn der Klassifikation für alle Datenpunkte mit Fehler.

Wie bei der Delegation ist auch hier die Problematik auf der taktischen Karte aufgetreten, bei der die Blickbewegungsmessung keine verlässlichen Aussagen über die wahrgenommenen Informationen geben konnte. Die Abschätzung basierend auf der Interaktion war in diesem Versuchsszenario zwar gut möglich, es stellt sich jedoch die Frage, inwiefern Aussagen bei Überwachungsaufgaben ohne Interaktionen auf einer Karte mit hoher Informationsdichte getroffen werden können.

### 4.3.3 DISKUSSION

Die Ergebnisse der beiden Versuche deuten auf einen Zusammenhang zwischen den Leistungsmaßen und der SB-Abweichung in aufgabenrelevanten **SB-KNOTEN** des **SB-MODELLS** hin. Nach Ausschluss von Ausreißern konnten mittlere bis starke Korrelationen in den individuellen Mittelwerten nachgewiesen werden und auch der subjektive Eindruck der Teilnehmer stand in einem Zusammenhang mit den Abweichungsmesswerten, was sich in der negativen Korrelation von Abweichung und der SART-Fragebögen äußerte.

Ein interessanter Aspekt des ersten Versuchs war, dass bei den Teilnehmern mit besonders schlechter Leistung in der Tiefflugaufgabe, der Fehler am stärksten mit hoher SB-Abweichung zusammenhing. Bei Teilnehmern im besseren Leistungsbereich war dieser Zusammenhang nicht so stark ausgeprägt (vgl. Abbildung 4-20). Es ist vorstellbar, dass hier eine untere Grenze der mittleren Abweichung gibt, die trotzdem ausreicht, um die Tiefflugaufgabe optimal auszuführen. Wie bereits diskutiert, muss die Radarhöhe nicht durchgängig überwacht werden, um trotzdem gute Leistungen zu erzielen. Der erste Versuch zeigte in diesem Kontext auch, dass die Probanden in einem Großteil der Fehlersituationen das Bewusstsein über die fehlerhaften Systemzustände hatten. Das ist nachvollziehbar, da die Korrektur eines Fehlers beispielsweise beim Tiefflug nicht instantan möglich ist. Auf der anderen Seite gilt es zu beachten, dass niedriges SB in einer Information nicht immer auf einen Fehler hindeutet. Dies gilt besonders in realen Aufgabenumgebungen, in denen die temporäre Konzentration auf eine Aufgabe automatisch dazu führt, dass andere Informationen vernachlässigt werden.

Die Untersuchung der Klassifikationsleistung zeigte, dass die SB-Abweichung in vielen Fällen den Grund für einen Fehler der Probanden aufdecken konnte, jedoch eine hohe SB-Abweichung nicht automatisch in Zusammenhang mit einem Fehler steht. Beim Bemerkern kritischer Informationen konnte die SB-Bewertung in einigen Situationen den Fehler des Piloten direkt messen, wie bei der Meldung der Flugabwehrstellungen oder beim Erkennen eines neuen Ziels gezeigt wurde. Hier zeigte

sich auch, dass die Verwendung von Interaktionen ein Sicherheitsmechanismus ist, der die fehlerhafte Blickbewegungsmessung ergänzen kann und falsch positive Ergebnisse verhindert. Dies verhindert, beispielweise, dass der Bediener ein Ziel laut SB-Bewertung nicht kennt, obwohl er schon längst damit interagiert hat.

Im zweiten Versuch konnte in einigen Fällen auch erkannt werden, wenn eine Information zu lange vernachlässigt wurde. Jedoch offenbarten sich hier zwei Probleme. Erstens war die Blickbewegungsmessung auf der taktischen Karte problematisch, da sich dort eine hohe Anzahl von Objekten auf kleiner Anzeigefläche befand. Durch die Auswertung von 2° um den gemessenen Fixationspunkt wurden hier also viele Informationen als wahrgenommen gewertet, obwohl die Probanden diesen wohl keine Aufmerksamkeit schenkten. Ein zweites Problem offenbarte sich in der nicht vorgesehenen Aufgabenausführung einiger Probanden. Diese wendeten Strategien zur Optimierung der Gesamtmissionsperformance an, bei denen vorsätzlich nicht so schnell wie möglich eine Auftragsdelegation durchgeführt wurde oder auch ein suboptimales UAV ausgewählt wurde. Dieser Fall zeigt aber auch, dass es nicht möglich ist, diese Art des „überblickenden“ SBs für Missionsziel und Dynamik der Situation mit dem **SB-MODELL** abzubilden. Mit Blick auf diese komplexen Entscheidungen der Probanden ist das **SB-MODELL** beschränkt auf die Erkennung von übersehenen Informationen.

Die Versuche konnten ebenfalls demonstrieren, dass die Verarbeitung einer **IDENTIFIKATION** in den Knoten sinnvoll ist, da vor allem die Informationen in Missionsaufgaben durch gleiche Objekte von unterschiedlicher Identität beschrieben werden konnten. So konnten im ersten Versuch beispielsweise die aktuelle Bedrohung und das Wissen über die spezifische Bedrohung erkannt werden. Diese Kontextinformation ist nützlich zur Verwendung in einem Assistenzsystem, da diese als aufgabenspezifische Informationen verwendet werden können.

### Limitierung der Versuchsgestaltung

Der Versuch hatte zwei Störfaktoren, die die Validität des Versuchs beeinflussen könnten. Zunächst ist die Teilnehmergruppe sehr heterogen, da erfahrene Kampfflugzeugpiloten mit Pilotenanwärtern zusammen ausgewertet wurden. Dieser Störfaktor könnte aber auch ein Qualitätskriterium des SB-Messsystems sein, wenn es in der Lage ist, die Varianz innerhalb der Probandengruppe und die damit verbundene Varianz in der Leistung vorherzusagen.

Darüber hinaus war der zweite Versuch nicht optimal gestaltet, sodass einige Probanden eine Taktik verfolgten, die nicht im Sinne der Aufgabenstellung war. Einige Probanden warteten bei einem neuen Ziel ab, bis die UAVs weniger Aufträge im Missionsplan hatten, um auszunutzen, dass in der Wartezeit potenziell neue Ziele erscheinen. Dies war ein Störfaktor für die Messung der Delegationszeit.

### Zusammenfassung

Aus den Ergebnissen konnten folgende zentrale Erkenntnisse für das SB-Messsystem gewonnen werden:

- Bei richtiger Modellierung der Ziele des Bedieners und der dafür notwendigen Informationen existiert ein Zusammenhang zwischen mittlerer SB-Abweichung und Leistungsmaßen sowie der Selbsteinschätzung.
- Das Messsystem kann informationsspezifische Aussagen zu fehlendem oder falschem SB geben. Dies gilt besonders für Überwachungs- und Entdeckungsaufgaben.

- Die informations- und identifikationspezifische Formulierung des Modells ermöglicht aufgabenspezifische Bewertung des SBs, welche auch nützlich für ein Assistenzsystem ist.
- Komplexere Aufgaben wie Missionsplanung bergen die Gefahr, dass der Bediener Ziele verfolgt und die damit zusammenhängenden Informationen nicht im SB-Modell vorgesehen sind.
- Eine hohe Abweichung in einem **SB-KNOTEN** bedeutet nicht automatisch, dass der Bediener einen Fehler macht. In Mehrfachaufgabensituationen vernachlässigen Bediener bewusst Informationen, um andere Informationen zu verarbeiten, ohne dass dabei die Leistung sinkt.

## 4.4 INTERVENTIONSGENERIERUNG

Die Validierungsversuche haben gezeigt, dass mithilfe der SB-Bewertung Lücken im SB eines Bedieners gefunden werden können. Die folgende Studie beschäftigt sich mit der Generierung von Intervention auf Basis dieser SB-Bewertung. Die Funktionsweise wurde bereits im Abschnitt 3.4 konzeptionell beschrieben und deren Implementierung in 3.5.2.5 vorgestellt.

Die zentrale Frage ist, ob ein SB-adaptives System – unter der Annahme, dass die SB-Bewertung richtige Ergebnisse liefert – die Leistung eines Mensch-Maschine-Systems steigern kann und ob diese Art der Unterstützung von Bedienern überhaupt als hilfreich und nützlich wahrgenommen wird. Die Strategie, dem Bediener nicht bewusste Informationen mitzuteilen könnte je nach Situation nützlich oder störend sein. Somit sollten die Validierungsversuche die folgenden Forschungsfragen beantworten:

- Kann die, auf der SB-Bewertung basierende, adaptive Assistenz eine Leistungssteigerung des Mensch-Maschine-Systems erzielen?
- Wie wird ein solches adaptives System von einem Bediener in Hinblick auf dessen Nutzbarkeit bewertet?

### 4.4.1 GESTALTUNG DER STUDIE

Um zu gewährleisten, dass die SB-Abschätzung valide Ergebnisse liefert, wurden im Versuch vergleichbare Aufgabentypen mit denen der bereits vorgestellten Versuchen herangezogen. Die im Folgenden beschriebenen Versuche wurden im Anschluss an die Versuche der Validierungsstudie der Aufmerksamkeitsanalyse durchgeführt (siehe Abschnitt 4.2). Um die Ergebnisse dieser Studie als Vergleich für die folgende Studie zu nutzen, wurden die beiden Versuchsaufgaben an die Missionen angelehnt, welche in Abschnitt 4.2.1.1 beschrieben sind. Aus diesem Grund sind auch die Versuchsteilnehmer sind gleichen und in Tabelle 4-5 beschrieben.

#### 4.4.1.1 Versuchsgestaltung

Im Folgenden werden die zwei durchgeführten Missionen mit adaptiver Interventionsgenerierung beschrieben. Die Aufgabentypen der beiden Missionen sind bereits aus dem Abschnitt 3.2 bekannt.

#### Mission 1: Instrumentenflug mit Anweisungen unter Böeneinfluss

Die erste Mission wurde so gestaltet, dass sie mit der Mission aus 4.2.1.1 vergleichbar ist. Die Aufgaben waren die gleichen, wobei eine Anpassung der Böenzeiten sowie der Positionierung der Wegpunkte vorgenommen wurde, welche in Abbildung 4-35 dargestellt ist. Die Zeiten der Anweisung sind die gleichen wie in Tabelle 4-2 dargestellt. Die Zeiten der Böen sind in der Appendix Tabelle 6 - Detaillierter Ablauf der Mission hinterlegt.

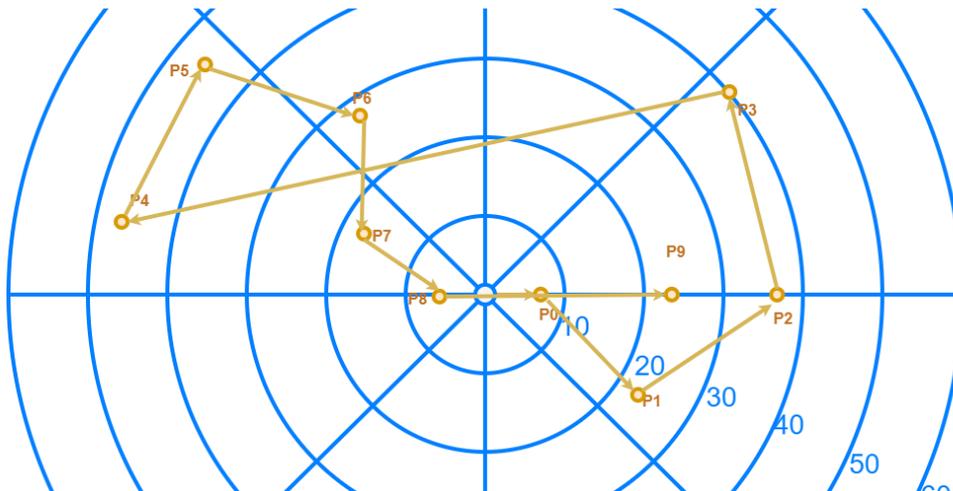


Abbildung 4-35 - Wegpunkte in der Mission 1

### Mission 2: Führung vonUCAVs

Im Gegensatz zur Mission im Kontext der Aufmerksamkeitsmessung wurde hier auf die Flugaufgabe verzichtet. So hatten die Versuchspersonen in dieser Mission nur die Aufgabe derUCAV-Führung zur Aufklärung oder Bekämpfung der auftauchenden taktischen Objekte. Die Abbildung 4-36 zeigt die Platzierung der Objekte im Missionsgebiet und die Appendix Tabelle 6 - Detaillierter Ablauf der Mission beschreibt den genauen zeitlichen Ablauf.

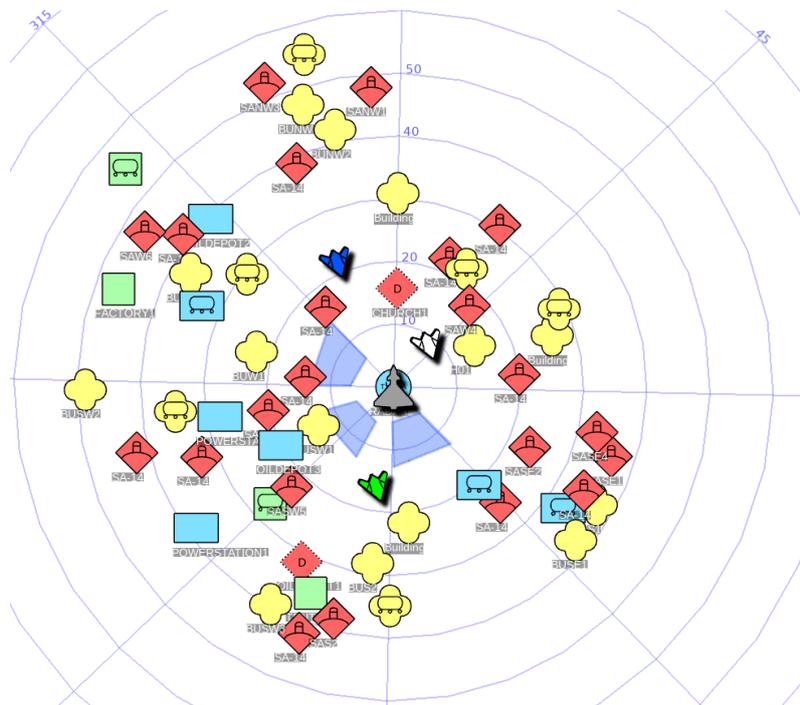


Abbildung 4-36 - Übersicht über die Lage in der Mission 2

### SB-Modelle

Das Modell auf dessen Basis das SB für den ersten Versuch abgeschätzt wird, ist in Abbildung 4-37 dargestellt. Das Modell beinhaltet alle relevanten Größen zur Bearbeitung der Aufgaben. Ein bemerkenswerter Aspekt liegt in der Modellierung des **SB-KNOTENS** „Flugpfadabweichung“. Die Flugpfadabweichung wird in der taktischen Karte nicht direkt visualisiert, sondern ist für den Bediener

über den Abstand zwischen Flugzeugsymbol und Flugpfad abzuschätzen, siehe Abbildung 4-38. Hier stellt sich die Frage, wie diese Abschätzung im **SB-MODELL** repräsentiert werden sollte. Eine Möglichkeit ist es, den Flugpfad mit dem **INHALT** Flugpfadkurs und den Flugzeugkurs getrennt als **SB-KNOTEN** zu modellieren und dann über eine **AND-RELATION** zu verknüpfen. Ändert sich nun eins dieser beiden **INHALTE** ohne, dass der Bediener dies wahrnimmt, kann eine SB-Abweichung festgestellt werden. Diese Modellierung hat sich in Tests als unvorteilhaft herausgestellt, da die Bediener nicht explizit Flugzeug und Flugpfad fixierten mussten, um die Information wahrnehmen zu können. Ein Blick in die Peripherie des Flugpfads reicht aus, um die Information wahrnehmen zu können. Um dieses messtechnische Problem zu umgehen, wurde der **SB-KNOTEN** Flugpfadabweichung erstellt und assoziiert mit einer Observation, die bei einer Fixation auf Flugpfad oder Flugzeugposition mit dem **INHALT** der räumlichen Distanz zwischen Flugzeug und Flugpfad generiert wurde.

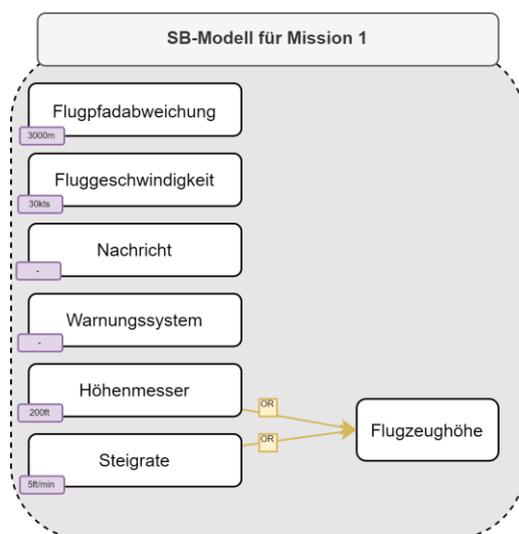


Abbildung 4-37 - SB Modell des ersten Versuchs. Der Parameter  $c_n$  ist in lila markiert.

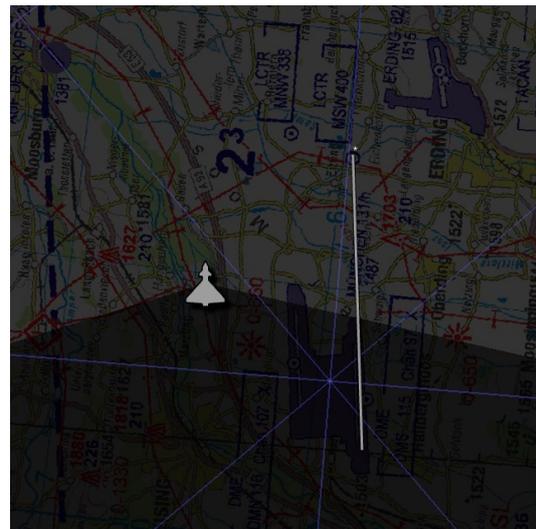


Abbildung 4-38 - Darstellung der Flugpfadabweichung

Die Abbildung 4-39 zeigt das **SB-MODELL** für den zweiten Versuch, welches sehr simpel ist und sich nur auf die Position neuer Objekte auf der taktischen Karte bezieht. Die Abweichungsberechnung ist in diesem Zusammenhang nicht relevant, da dieses Modell nur aussagen kann, ob ein neu aufgetauchtes Objekt bereits vom Bediener wahrgenommen wurde. Im Versuch bewegen sich auch die Fahrzeuge nicht, somit gibt es bei den taktischen Objekten keine dynamische Veränderung.

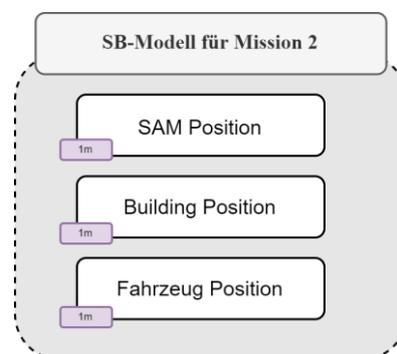


Abbildung 4-39 - SB Modell des zweiten Versuchs.

## Interventionen

Es wurden zwei verschiedene Arten von Interventionen umgesetzt: Aufmerksamkeitslenkung auf der taktischen Karte und auditive Textbenachrichtigungen basierend auf einer Sprachsynthese-Software.

### Aufmerksamkeitslenkung auf der taktischen Karte

Um die Aufmerksamkeit auf taktische Objekte zu lenken, wurden farblich hervorgehobene Kreise mit einer Textannotation generiert. Diese Benachrichtigungen, dargestellt in Abbildung 4-40, blieben acht Sekunden lang sichtbar.



Abbildung 4-40 - Hervorhebung von taktischen Objekten: Gebäude (links) und Flugabwehrstellung (rechts)

Für den Fall, dass das taktische Objekt nicht im aktuellen Ausschnitt der Karte liegt, wurde ein blinkender Pfeil implementiert, der auf eine Hervorhebung außerhalb des Kartenausschnitts aufmerksam machen sollte (siehe Abbildung 4-41).

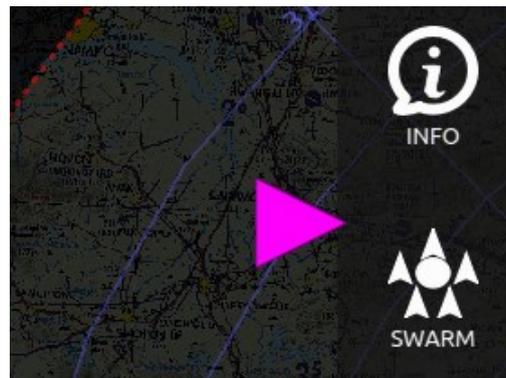


Abbildung 4-41 - Pfeil am Rand der taktischen Karte im MHDD, der die Richtung des hervorgehobenen Objekts visualisiert.

### Auditive Textnachrichten

Die Benachrichtigungen waren so gestaltet, dass sie nur die Systemgröße benannten, die gerade vom Bediener vernachlässigt wurde. In diesem Fall waren die Benachrichtigungen „ALTITUDE“, „VELOCITY“, „ROUTE“, „MESSAGE“ und „WARNING“.

### Parametrisierung der Assistenzfälle

Der im Abschnitt 3.5.2.5 vorgestellte Algorithmus verlangt für jeden Assistenzfall Zeitkonstanten (siehe *Time-Until-Notification* in Algorithmus 3-2), welche die Zeit zwischen hoher gemessener Abweichung und dem Auslösen einer Intervention definieren. Die Tabelle 4-9 und Tabelle 4-10 zeigen die interventionsspezifischen Parameter.

Assistenz	Assoziierter Knoten	Time-until-notification [s]
„ALTITUDE“	Flugzeughöhe	2
„VELOCITY“	Fluggeschwindigkeit	2
„MESSAGE“	Nachricht	5
„WARNING“	Warnsystem	5
„ROUTE“	Flugpfadabweichung	3

Tabelle 4-9 - Parametrisierung der Interventionsgenerierung für Versuch 1

Die Parameter wurden mittels vorher durchgeführter Probeläufe gewählt und waren abgestimmt auf die Dynamik der Systemgrößen sowie die Frequenz der Störungen.

Assistenz	Assoziierter Knoten	Time-until-notification [s]
„NEW BUILDING“	Building Position	5
„NEW VEHICLE“	Fahrzeug Position	5
„GROUND THREAT“	SAM-Site Position	5

Tabelle 4-10 - Parametrisierung der Interventionsgenerierung für Versuch 2

#### 4.4.1.2 Messgrößen & Datenverarbeitung

Während des Versuchs wurden Daten zur SB-Abweichung, Aufgabenleistung, Interventionsanzahl und Blickbewegung aufgezeichnet. Darüber hinaus wurde eine subjektive Bewertung des Assistenzsystems eingeholt.

#### Leistungsmaße

Die Aufgabenleistung für die Verfolgungsaufgaben wurde wie folgt quantifiziert: Die mittlere Fehlerrate wurde als quadratisches Mittel der Differenz zwischen dem Soll- und dem Ist-Wert berechnet. Wenn der tatsächliche Wert unter einem Grenzwert lag, wurde der Fehler auf  $e = 0$  gesetzt. Die vorgegebenen Grenzwerte für die Verfolgung waren  $\Delta e_{\text{Flughöhe}} = 200 \text{ft}$ ,  $\Delta e_{\text{Geschw.}} = 15 \text{kt}$  und  $\Delta e_{\text{Flugpfad}} = 2 \text{Nm}$ .

Um den Einfluss des Systems auf die Verteilung der visuellen Aufmerksamkeit in allen Aufgaben zu quantifizieren, wird die Zeit zwischen Aktivierung einer Aufgabe - beispielsweise eine Windbö oder eine neue Nachricht - und der ersten Fixation auf die relevante Anzeige im Cockpit gemessen. Diese Erkennungsdauer war also definiert als:

$$T_{\text{Erkennung}}^{\text{Aufgabe}} = T_{\text{Fixation}}^{\text{Aufgabe,AoI}} - T_{\text{Aktiv}}^{\text{Aufgabe}} \quad (10)$$

Als Leistungsmaß für die Mission 2 wurde die Delegationszeit zwischen Erscheinen eines Ziels auf der taktischen Karte und erster Interaktion wie folgt definiert:

$$T_{\text{Delegation}}^{\text{Ziel}} = T_{\text{Interaktion}}^{\text{Ziel}} - T_{\text{Erscheinen}}^{\text{Ziel}} \quad (11)$$

#### Fragebogen

Zur Bewertung der Nutzbarkeit wurde ein Fragebogen erstellt, der nach jeder Mission ausgefüllt wurde. Der Fragebogen ist im Anhang E.6.3 dokumentiert. Die ersten drei Fragen bezogen sich auf Zeitpunkt und Frequenz der Meldung, sowie Nachvollziehbarkeit der Zeitpunkte. Die Fragen 4-8

sollten klären, inwiefern die Interventionen den normalen Arbeitsfluss der Piloten unterstützten oder störten. Das Vertrauen in das System wurde in Frage 9 und 10 abgefragt und Frage 11 war eine Zusammenfassung, ob das System als hilfreich eingestuft wird. Die darauffolgenden Fragen sollten im Vergleich zum Versuch ohne Assistenzsystem getroffen werden. Hierbei ginge es einerseits um die Einschätzung, ob sich die Leistung verbessert und die Belastung oder Frustration vergrößert hat. Außerdem wurde abgefragt, ob die Piloten bewusst ihr Verhalten geändert haben, indem sie sich beispielsweise auf Meldungen verlassen haben. Im zweiten Teil sollten dem Assistenzsystem Attribute in den Dimensionen Verständlichkeit, Nachvollziehbarkeit, Vorhersehbarkeit, Zuverlässigkeit, Aufdringlichkeit und Intelligenz zugeordnet werden.

### Datenverarbeitung

Um die Normalverteilung der Daten zu prüfen, wurde der Shapiro-Wilk-Test verwendet. Wenn der Test eine Normalverteilung anzeigte, wurde der Korrelationskoeffizient von Pearson und ein abhängiger t-Test als Signifikanztest verwendet. Andernfalls wurde der Rangkorrelationskoeffizient von Spearman und der Wilcoxon-Signed-Rank-Test verwendet. Für die Gleichheit der Varianzen wurde der Levene-Test verwendet (Gastwirth et al., 2009). Der Effekt wurde als schwach eingestuft, wenn  $r \leq 0.1$ , mittel, wenn  $0.1 < r \leq 0.7$  und stark, wenn  $r > 0.7$ . Das Signifikanzniveau wurde auf  $p < 0.05$  festgelegt.

#### 4.4.1.3 Ablauf

Die beiden vorliegenden Versuche wurden in Anschluss an die in Abschnitt 4.2.1.4 beschriebenen Versuche durchgeführt. Daher kannten die Probanden die Aufgaben bereits und wurden lediglich darüber in Kenntnis gesetzt, dass sie in den folgenden Versuchen durch ein Assistenzsystem unterstützt werden. Die Funktionsweise des Systems wurde nicht erklärt, um ein unvoreingenommenes Meinungsbild zur Funktionalität einholen zu können. Die Probanden füllten nach jedem Durchgang den Fragebogen aus. Im Anschluss an beide Versuche gab es noch eine Nachbesprechung, in der die Piloten weitere Rückmeldung und Verbesserungsvorschläge über das System geben konnten.

### 4.4.2 ERGEBNISSE

Um einen Vergleich der Leistung zwischen aktiver und inaktiver Interventionsgenerierung zu ermöglichen, wurde der Versuch aus Abschnitt 4.2 nochmals herangezogen und um die Bewertung der SB mithilfe des eben vorgestellten Modells ergänzt. Diese Ergebnisse werden im Folgenden für den Vergleich verwendet.

#### 4.4.2.1 Mission 1

##### Validierung der SB-Messung

Um sicherzustellen, dass sich auch in diesem Versuch ein Zusammenhang zwischen der SB-Messung und den Leistungsmaßen existiert, wurde zunächst die Fehlerraten aus dem ersten Versuch ohne Assistenzsystem mit den gemessenen SB-Abweichungswerten verglichen. Die Abbildung 4-42 zeigt die durchschnittlichen Werte pro Probanden.

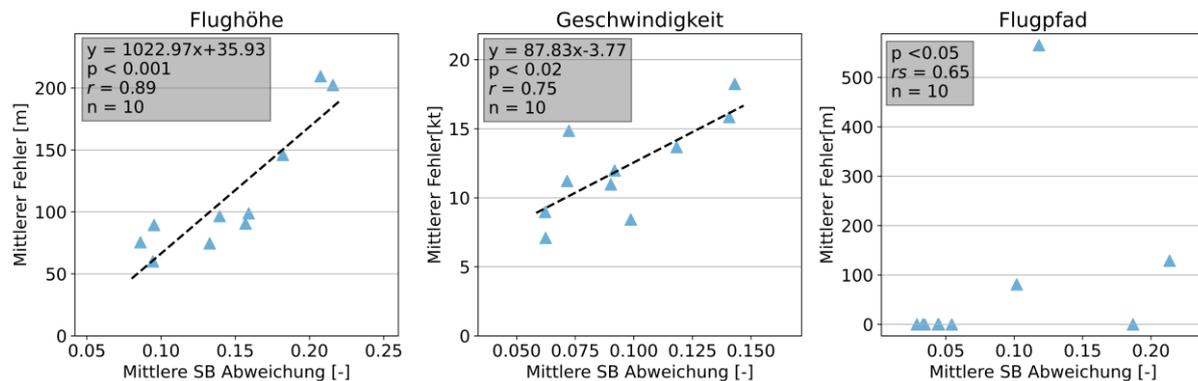


Abbildung 4-42 - Zusammenhang zwischen mittleren Fehlern in den drei Verfolgungsaufgaben Flughöhe, Geschwindigkeit sowie Flugpfad und der SB-Abweichung in der damit verbundenen, relevanten Systeminformation

Die Fehler in den Aufgaben Flughöhe und Geschwindigkeit zeigten signifikante positive Korrelationen mit der SB-Abweichung. Der Korrelationskoeffizient zwischen Flughöhe und SB-Abweichung im Knoten „Flughöhe“ betrug  $r = 0.89$ . Eine etwas schwächere Korrelation zeigt sich bei der Geschwindigkeit und der Abweichung mit dem Koeffizienten von  $r = 0.75$ . Der Zusammenhang zwischen Fehler beim Flugpfad und der SB-Abweichung ist weniger gut nachzuweisen, da sechs der zehn Probanden keine Fehler in dieser Aufgabe begangen haben. Der Spearman Korrelationskoeffizient dieser nicht normalverteilten Daten betrug  $r_s = 0.65$ . Die Korrelationen waren in diesem Versuch sogar stärker ausgeprägt als im Validierungsversuch aus Abschnitt 4.3, was durch die Unvorhersehbarkeit der Böen erklärt werden könnte. Beim Tiefflugszenario konnten Änderungen in der Höhe durch den Terrainverlauf antizipiert werden. In diesem Szenario war das durch die künstliche Störung des Flugzeugs zu unbekanntem Zeitpunkten nicht möglich.

### Zusammenhang zwischen Fehler und Interventionen

Die Interventionen wurden durch hohe Abweichungen in der gemessenen SB-Abweichung ausgelöst. Die Abbildung 4-43 zeigt die Anzahl an Interventionen im Versuch mit aktivem Assistenzsystem gegenüber dem mittleren Fehler getrennt nach den drei kontinuierlichen Aufgaben.

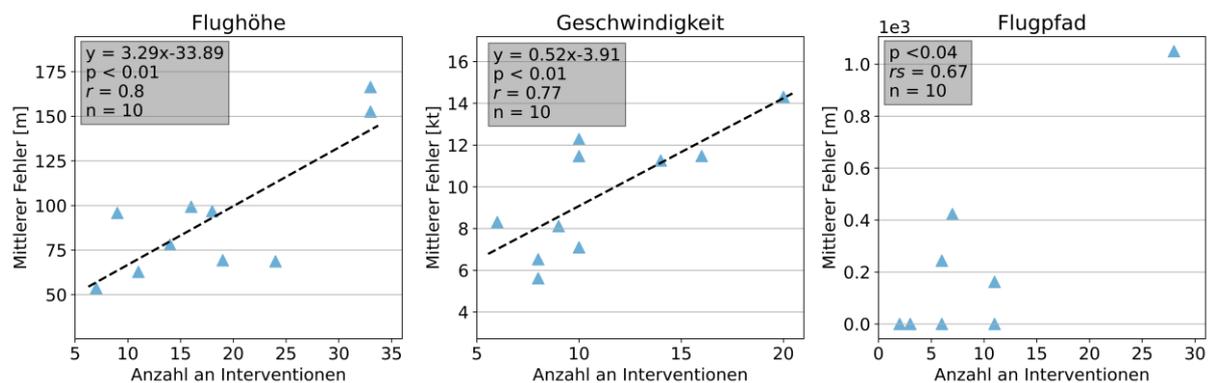


Abbildung 4-43 - Zusammenhang zwischen Anzahl an Interventionen und der SB-Abweichung in den drei Verfolgungsaufgaben Flughöhe, Geschwindigkeit und Flugpfad

Hier zeigt sich, dass Piloten mit hohem mittlerem Fehler auch viele Interventionen des Assistenzsystems bekamen. Zwischen mittlerer Fehlerrate und der Interventionsanzahl in den Aufgaben Flughöhe und Geschwindigkeit gibt es mit  $r = 0.8$  und  $r = 0.77$  stark positive

Korrelationen, die statistisch signifikant sind. Eine Besonderheit des Interventionsmechanismus zeigt sich in der Anzahl der Interventionen beim Flugpfad. Diese wurden auch bei denjenigen Probanden ausgelöst, welche keine Fehler begangen haben. Die Interventionsgenerierung hat in diesem Versuch lediglich überprüft, ob die Abweichung hoch ist und nicht, ob tatsächlich ein Fehler vorliegt. Wenn dieser für einen längeren Zeitraum zu hoch war, wurde eine Benachrichtigung veranlasst, selbst wenn kein Fehler vorlag.

In diesen Daten zeigt sich die Tendenz der adaptiven Strategie, Bedienern bei vielen Fehlern auch viele Hinweise zu geben. Dies muss kritisch begutachtet werden, da eine erhöhte Zahl der Interventionen auch die Arbeitsbelastung und damit die Beanspruchung des Bediener erhöht. Aus dieser Erhöhung resultiert möglicherweise eine geringere Leistung, was zu weiteren Interventionen führt.

### Vergleich der Leistung im Instrumentenflug

Nach der Identifikation grundsätzlicher Zusammenhänge, wird sich im Folgenden dem Vergleich zwischen der Leistung in der Mission mit gegenüber der ohne Assistenzsystem gewidmet. Hierfür zeigt die Abbildung 4-44 einen Vergleich der mittleren Fehler in beiden Versuchen.

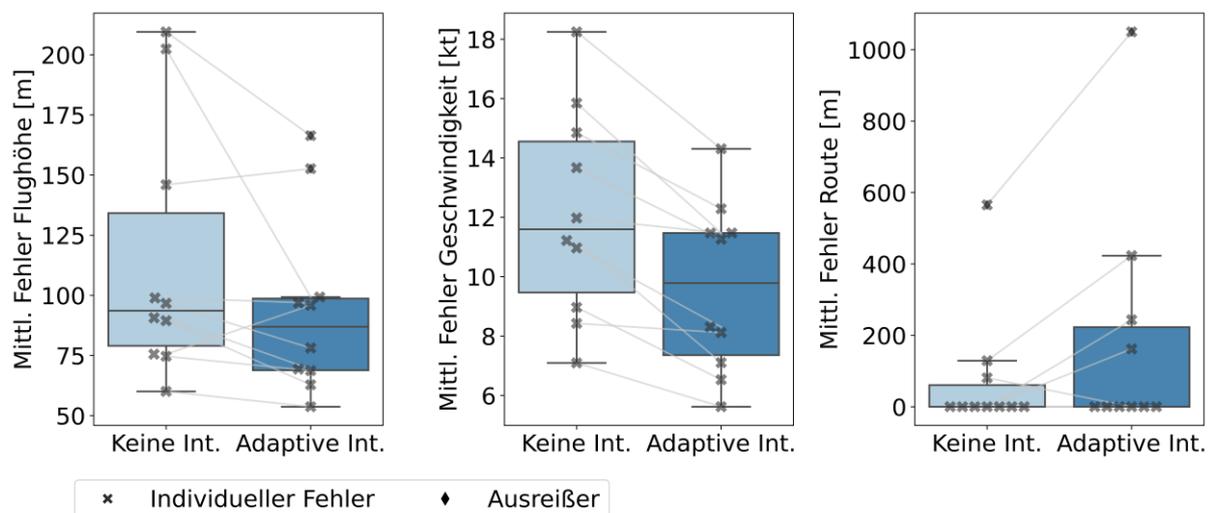


Abbildung 4-44 - Vergleich der Leistung in den drei Verfolgungsaufgaben zwischen der Mission mit und ohne SB-adaptive Intervention. Die Kreuze stehen für einen Probanden und sind über die zwei Missionen miteinander verbunden, um die individuelle Veränderung der Leistung pro Probanden zu visualisieren.

Der Fehler in der Flughöhe und Geschwindigkeit reduzierte sich mit Assistenzsystem, wobei die Reduktion in der Flughöhe nicht signifikant war ( $p_s < 0.08$ ). Die Reduktion in der Geschwindigkeit ist nach einem gepaarten t-Test signifikant ( $t(9) = 5.53, p = 0.0004$ ). Gegenüber der Verbesserung in diesen beiden Aufgaben steht eine nicht signifikante Verschlechterung ( $p_s < 0.11$ ) der Leistung in der Flugpfadaufgabe. Eine Erklärung für die Verschlechterung der Leistung in dieser Aufgabe könnte sein, dass die Piloten durch die Interventionen in Bezug auf die anderen Aufgaben abgelenkt wurden und in den entscheidenden Momenten zur Verfolgung eines neuen Kurses vom Assistenzsystem unterbrochen worden sind. Die Abbildung 4-43 zeigt auch, dass sehr viel mehr Interventionen für Höhe und Geschwindigkeit generiert wurden, was der Volatilität dieser Systemgrößen durch natürliche Schwankung und Störung entspricht.

Die dargestellte durchschnittliche Leistung spiegelt zwei Aspekte der Aufgabenerfüllung wider. Erstens die Fähigkeit der Teilnehmer, die Verfolgungsaufgaben ohne Störung der Höhe und Geschwindigkeit des Flugzeugs auszuführen. Zweitens, die Fähigkeit, eine Störung schnell zu

bemerken und zum Zielwert zurückzukehren. Der Unterschied zwischen den beiden Bedingungen kann genauer analysiert werden, indem die Zeit wird, bis die Teilnehmer einen Fehler im Zustand des Flugzeugs bemerken, nachdem er herbeigeführt wurde. Dies soll im folgenden Abschnitt geschehen.

### Erkennungsleistung

Die Erkennungsleistung ist definiert über die Zeit  $T_{Erkennung}^{Aufgabe}$ , die die Dauer angibt, nachdem die Änderung einer relevanten Systemgröße durch eine versuchsseitige Störung das erste Mal fixiert wurde. Die Tabelle 4-11 gibt den Durchschnitt und die Standardabweichung der Erkennungsdauern aller Piloten an. In allen drei Aufgaben verbesserte sich die mittlere Erkennungsdauer und die Varianz im Durchgang mit der Intervention auf SB-Basis.

Bedingung	Höhenänderung [s]		Geschwindigkeitsänderung [s]		Flugpfadänderung [s]	
	M	SD	M	SD	M	SD
<b>Keine Intervention</b>	0.93	1.04	1.25	1.56	1.01	1.84
<b>Adaptive Intervention</b>	0.83	0.75	1.09	1.19	1.70	2.74

Tabelle 4-11 - Durchschnitt und Standardabweichung der Erkennungsdauern für die Verfolgungsaufgaben

Die Verteilung der Erkennungszeiten ist in Abbildung 4-45 dargestellt und zeigt, dass in den Aufgaben Flughöhe und -geschwindigkeit die großen Ausreißer reduziert wurden. Die Verringerung der Standardabweichung ist in beiden Fällen signifikant (Flughöhe  $p < 0.01$ , Fluggeschwindigkeit  $p < 0.03$ ). Hier konnte das adaptive System also verhindern, dass eine Systemgröße lange vernachlässigt wird, wobei das System frühestens nach zwei Sekunden eine Intervention auslöste. In der Flugpfadaufgabe sieht man einen gegensätzlichen Effekt mit einer höheren Standardabweichung aber niedriger mittlerer Erkennungsdauer. Die Veränderung der Standardabweichung war in diesem Fall nicht signifikant, wobei hier auch weniger Datenpunkte produziert wurden.

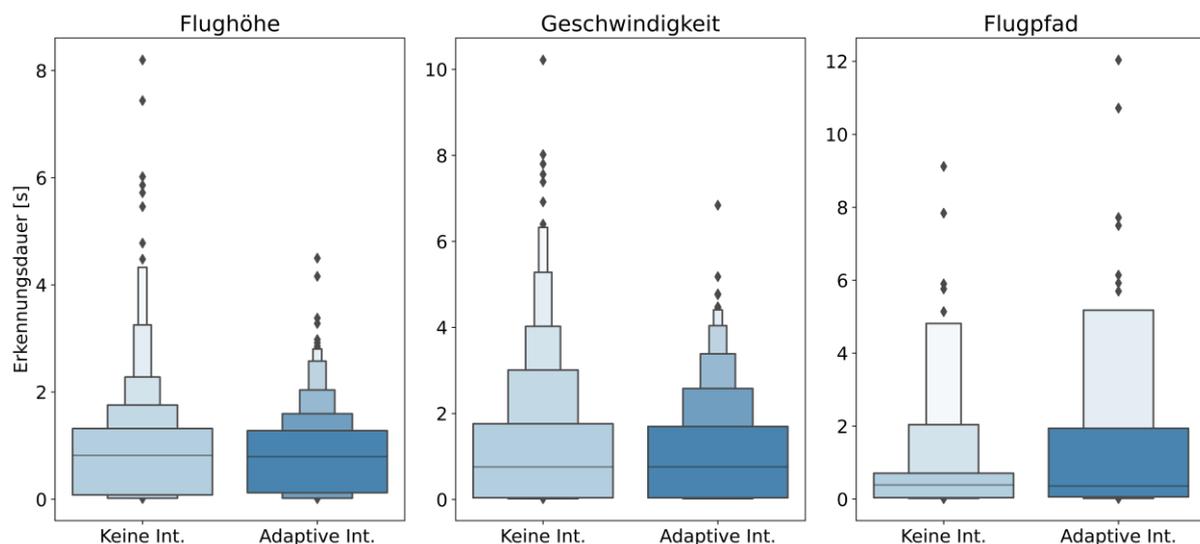


Abbildung 4-45 - Erkennungszeiten in den drei Aufgaben Flughöhe, Geschwindigkeit und Flugpfad

Analog hierzu stellt die Tabelle 4-12 die Daten für den Durchschnitt für die Erkennung einer Warnung sowie einer neuen Nachricht dar.

Bedingung	Warnung [s]		Nachricht [s]	
	M	SD	M	SD
Keine Intervention	3.96	11.86	7.93	22.49
Adaptive Intervention	1.23	.94	3.18	3.45

Tabelle 4-12 - Warnungen ( $N_{\text{Warnung}} = 20, N_{\text{Nachricht}} = 54$ )

Die Änderung der Werte zwischen den beiden Gruppen kann auf die Reduktion einzelner Ausreißer zurückgeführt werden, was auch in der Abbildung 4-46 dargestellt ist. Die Anzahl der Datenpunkte ist hier sehr gering, da nur zwei Warnungen und sechs Nachrichten pro Versuch vorkamen.

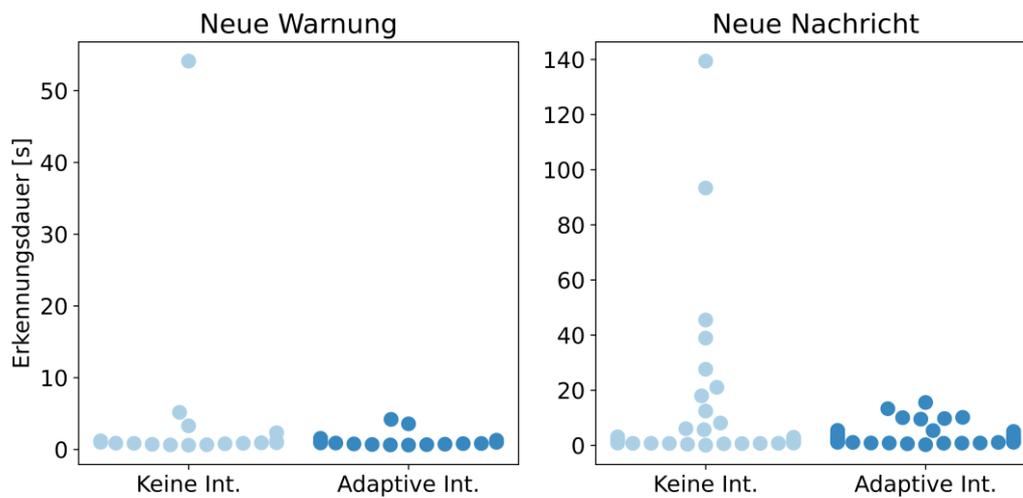


Abbildung 4-46 - Erkennungsdauern in den Aufgaben Warnung und Nachricht

### Subjektive Bewertung

Im Folgenden sollen die Ergebnisse des Fragebogens vorgestellt werden. Die Abbildungen zeigen farblich kodiert die Antworten auf der Bewertungsskala von 1 bis 8 (1 = „Trifft voll und ganz zu“, 8 = „Trifft überhaupt nicht zu“). Auf der horizontalen Achse sind die zehn Probanden in positive und negative Richtung aufgetragen, sodass die Anzahl der Antworten direkt abgelesen werden kann.

Die in Abbildung 4-47 dargestellte Bewertung in Bezug auf den Zeitpunkt und Nachvollziehbarkeit der Intervention ist sehr heterogen. Diese Punkte wurden tendenziell negativ bewertet, wobei zwei Piloten die beste Bewertung vergaben. Die Frequenz wurde dagegen als angemessen bewertet.

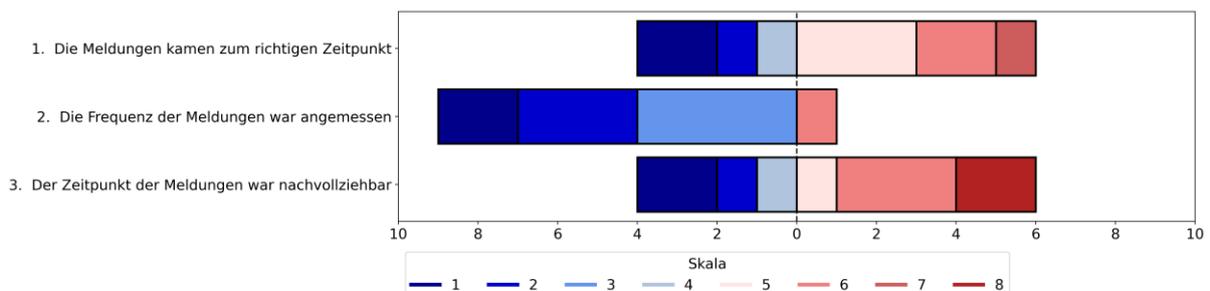


Abbildung 4-47 - Bewertung der zeitlichen Aspekte

Die Ergebnisse in Bezug auf den Arbeitsfluss und potenzielle Störung durch die Interventionen sind in Abbildung 4-48 dargestellt. Hier ergibt sich ein durchweg positives Bild, da die Meldung als überwiegend nicht aufdringlich oder störend betrachtet wurden. Die Interventionen zogen Aufmerksamkeit auf sich, ohne den Arbeitsablauf stark zu stören, wobei es hier einen Probanden gab, der diese positive Einschätzung nicht teilte.

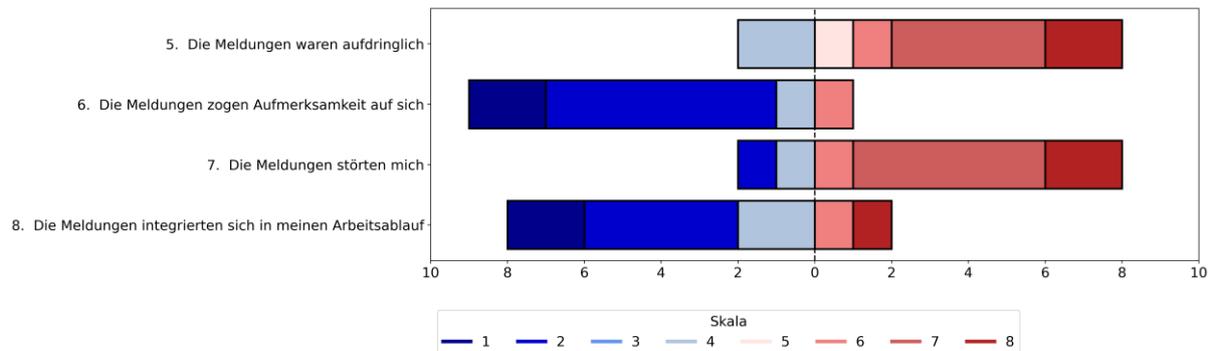


Abbildung 4-48 - Bewertung der Interventionen in Bezug auf den Arbeitsfluss

Abbildung 4-49 visualisiert die Ergebnisse auf Fragen des Vertrauens und der allgemeinen Nützlichkeit des Systems. Die Meldungen wurden von den Probanden beachtet, wobei sich beim Vertrauen ein gemischtes Bild ergibt. Einige Probanden verließen sich auf das System, während andere dem System nicht trauten. Sieben Probanden bewerteten das System als nützlich und drei als nicht nützlich.

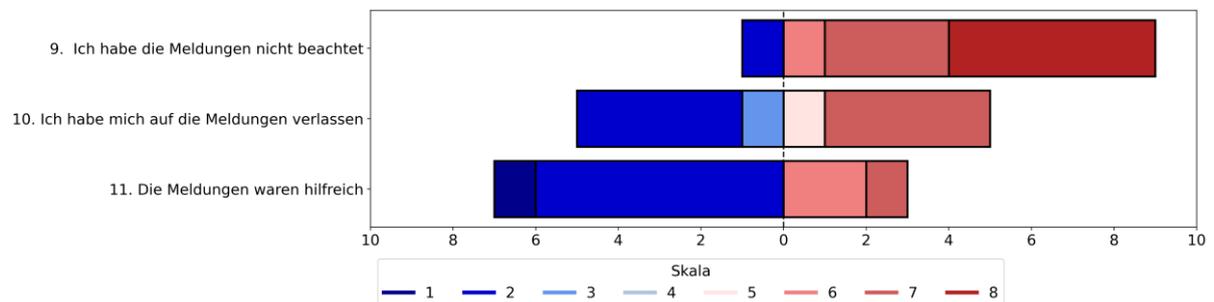


Abbildung 4-49 - Bewertung in Bezug auf Vertrauen und Nützlichkeit

Im Vergleich zum Versuch ohne Assistenz, waren 90% der Probanden überzeugt, dass sich ihre Leistung verbessert hat, wobei 20% auch höher belastet und 30% frustrierter waren (siehe hierzu Abbildung 4-50). Die Frage, ob sie ihr Verhalten im Versuch geändert haben, bejahten 7 von 10 Probanden.

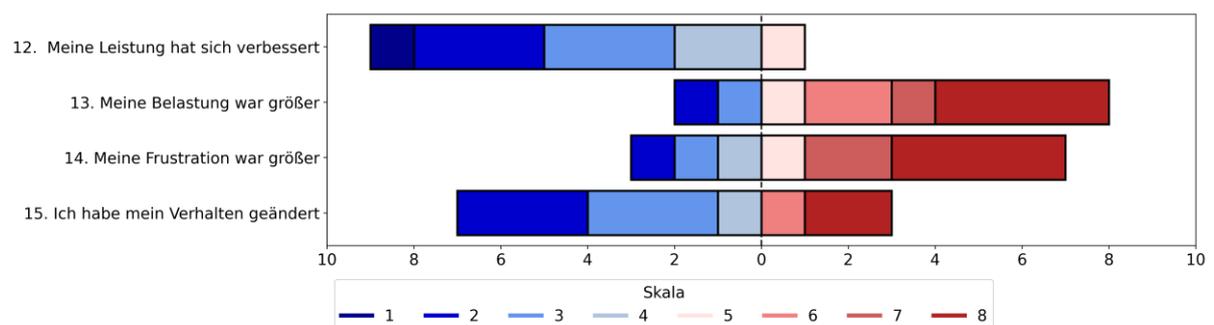


Abbildung 4-50 - Bewertung im Vergleich zum Versuch ohne Interventionen

Bei der Zuweisung von Attributen zeigt sich ebenfalls ein gemischtes Bild. Das Assistenzsystem wird tendenziell als verständlich, nachvollziehbar und vorhersehbar bewertet. Nach Besprechung der Antworten wurde „Verständlichkeit“ von den Probanden aber im Sinne der Verständlichkeit der Text-to-Speech-Benachrichtigungen gesehen. Ein sehr heterogenes Bild wird bei Zuverlässigkeit gezeichnet, was zu den vorhergehenden Bewertungen passt. Ferner wird das System als tendenziell vorschnell und dumm eingeschätzt.

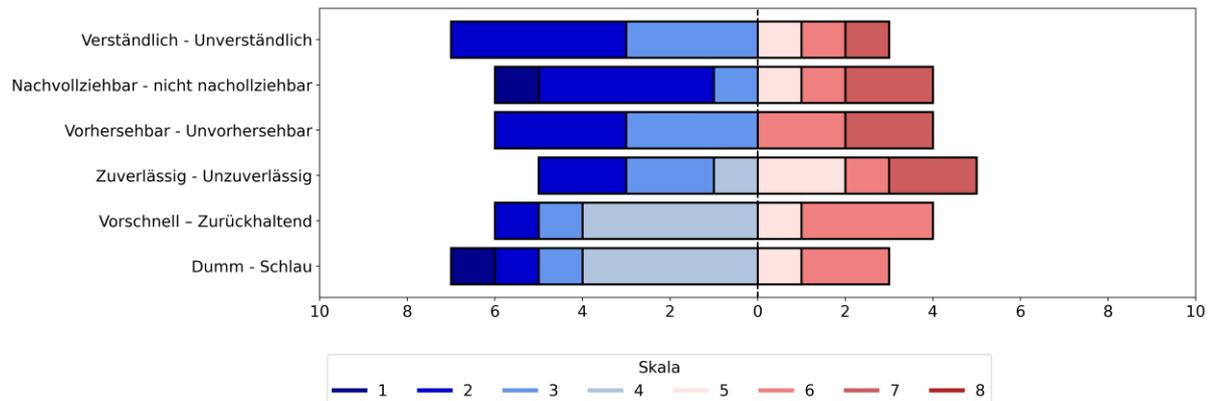


Abbildung 4-51 - Zuweisung von Attributen an die Interventionsgenerierung

Zusammenfassend hatte die subjektive Bewertung eine hohe Varianz unter den Versuchsteilnehmern mit einer Tendenz zur positiven Bewertung. Dies könnte damit erklärt werden, dass die Probanden unterschiedliche Leistung und Verhalten zeigten, was sich auch im adaptiven Verhalten der Interventionsgenerierung niederschlug. Ein Vergleich der Bewertung mit der Anzahl der Interventionen offenbarte jedoch keinen quantitativen Zusammenhang.

Bei den offenen Fragen im Fragebogen und bei der Diskussion mit den Probanden wurden folgende Punkte angebracht:

- Interventionen waren nicht immer notwendig, wenn eine Korrektur bereits eingeleitet wurde.
- Einige Probanden verloren das Vertrauen, nachdem Interventionen auch ausgelöst wurden, obwohl kein Fehler vorlag. Es gab die klare Erwartungshaltung, dass eine Intervention nur im Falle eines Fehlers mit Handlungsbedarf ausgelöst werden sollte. Besonders störend empfand ein Proband wiederholte falsch positive Interventionen.
- Auditive Warnung sollten nur für kritische Warnungen verwendet werden und nicht für minimale Abweichungen in der Höhe.
- Das Assistenzsystem war vorschnell und ließ einem Probanden nicht genug Zeit, die Aufgabe selbst zu erledigen.
- Ein Proband verließ sich auf das System und überwachte nicht mehr die Nachrichtenanzeige, sondern überließ es dem System.

#### 4.4.2.2 Mission 2

Die Ergebnisse der zweiten Mission werden im Folgenden vorgestellt. Primäre Aufgabe in diesem Versuch war eine möglichst schnelle Delegation von Aufträgen auf auftauchende Objekte. Bei Proband P1 traten Probleme beim Aufzeichnen der Daten auf. Beim Proband P10 trat ein Fehler bei der

Durchführung des Versuchs mit Interventionen auf, weshalb nur die ersten 30 Objekte auftauchten und somit nur diese in die Bewertung einfließen konnten.

### Zusammenhang zwischen Delegationszeit und Interventionen

Der Zusammenhang zwischen der Interventionsanzahl und der durchschnittlichen Delegationszeit ist in Abbildung 4-52 dargestellt. Beide Datensätze sind laut Shapiro-Wilk-Test nicht normalverteilt oder nur knapp an der definierten Grenze von  $p=0.05$ .

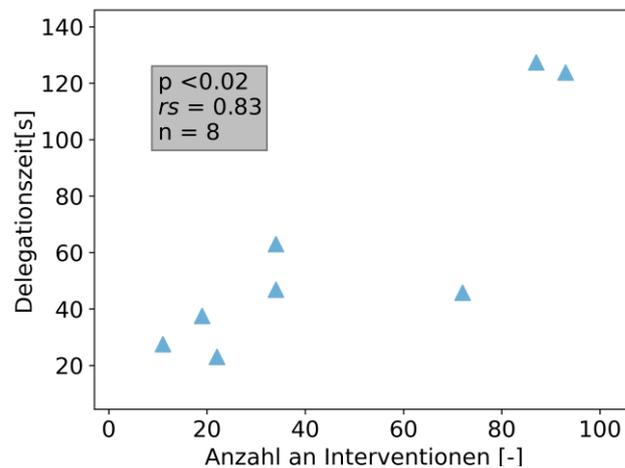


Abbildung 4-52 - Zusammenhang zwischen Interventionsanzahl und mittlerer Delegationszeit

Der Korrelationskoeffizient nach Spearman zeigt eine signifikant hohe Korrelation von  $r_s = 0.85$  an. Der statistische Trend der Daten zeigt auch, dass Probanden mit langer mittlerer Delegationszeit tendenziell mehr Benachrichtigungen bekamen, ähnlich zum Ergebnis des vorherigen Versuchs.

### Vergleich der Leistung in der Delegation

Beim Vergleich des Medians aller mittleren Leistungen, dargestellt in Abbildung 4-53, zeigt sich eine nicht-signifikante, geringe Verbesserung von  $T_{Del,Keine\ Int} = 48.78s$  auf  $T_{Del,Int} = 45.77s$ .

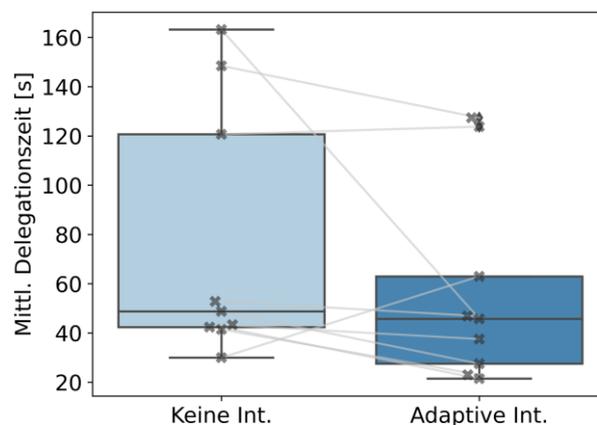


Abbildung 4-53 - Vergleich der mittleren Delegationszeit zwischen den Versuchen ohne und mit adaptiver Intervention

Ein Proband zeigt hier eine starke Verbesserung und ein anderer Proband eine massive Verschlechterung der mittleren Delegationszeit. Bei den restlichen Probanden konnte mit einer Ausnahme eine geringe Verbesserung gemessen werden.

### Erkennungsleistung

Die Tabelle 4-13 zeigt den Durchschnitt und die Standardabweichung der Erkennungszeiten für beide Bedingungen. Durch die Intervention kann die mittlere Erkennungszeit sowie deren Varianz reduziert werden.

Bedingung	Erkennungszeit [s]	
	M	SD
<b>Keine Intervention</b>	27.21	49.54
<b>Adaptive Intervention</b>	18.65	28.15

Tabelle 4-13 - Durchschnitt und Standardabweichung der Erkennungszeiten für die Delegationsaufgabe ( $N_{keine\ Int}=275$ ,  $N_{Int}=290$ )

Die unterschiedliche Anzahl an Datenpunkten kommt dadurch zustande, dass bei manchen Objekten keine Erkennung und auch keine Interaktion gemessen wurde, der Proband hat diese somit nicht bemerkt oder vergessen, einen Auftrag auszuführen. Abbildung 4-54 zeigt die Verteilung der Daten, in denen man die Reduzierung der Standardabweichung erkennen kann.

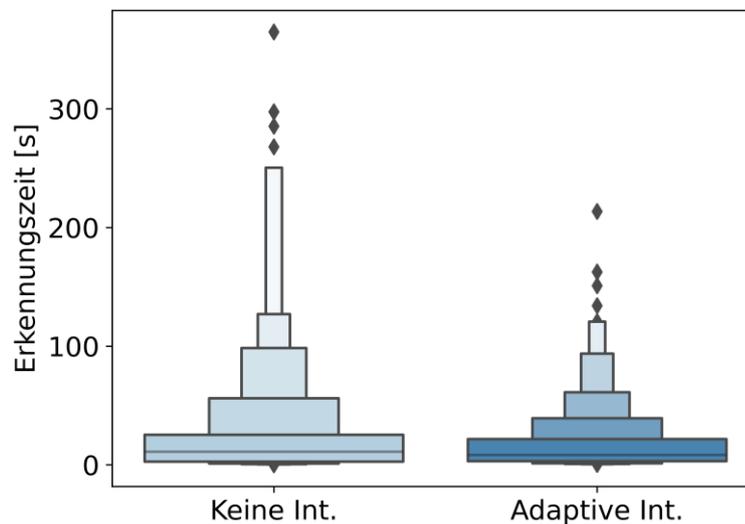


Abbildung 4-54 - Erkennungszeiten in der Delegationsaufgabe

### Subjektive Bewertung

Die Interventionen wurden von den Probanden sehr unterschiedlich bewertet. Beim Ausfüllen des Fragebogens wurde vom Probanden 8 vergessen, die ersten 15 Fragen zu beantworten. Aus diesem Grund gibt es dort nur neun Datenpunkte. Abbildung 4-55 zeigt die Bewertung der zeitlichen Aspekte, bei denen der Zeitpunkt und die Frequenz für über 60% der Probanden nachvollziehbar war.

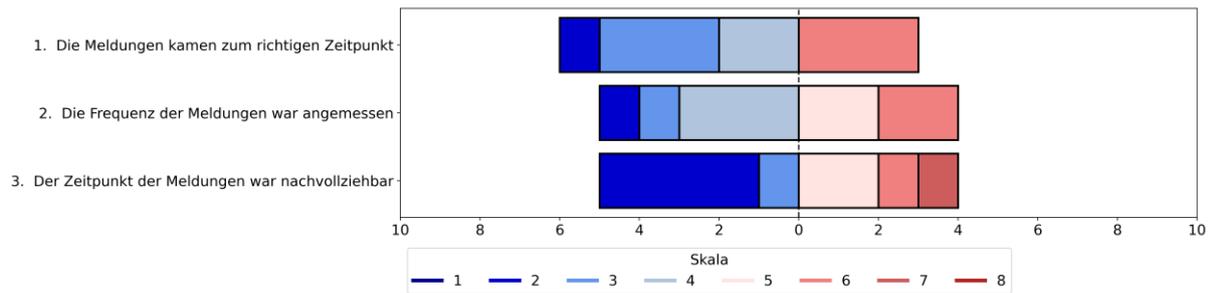


Abbildung 4-55 - Bewertung in Bezug auf zeitliche Aspekte

In Bezug auf den Arbeitsfluss, dargestellt in Abbildung 4-56, waren die Meldungen für 80%-90% der Probanden aufdringlich beziehungsweise zogen Aufmerksamkeit auf sich. Die Meldungen störten die Mehrheit der Probanden nicht, jedoch integrierten sie sich auch nicht bei allen gut in den Arbeitsablauf.

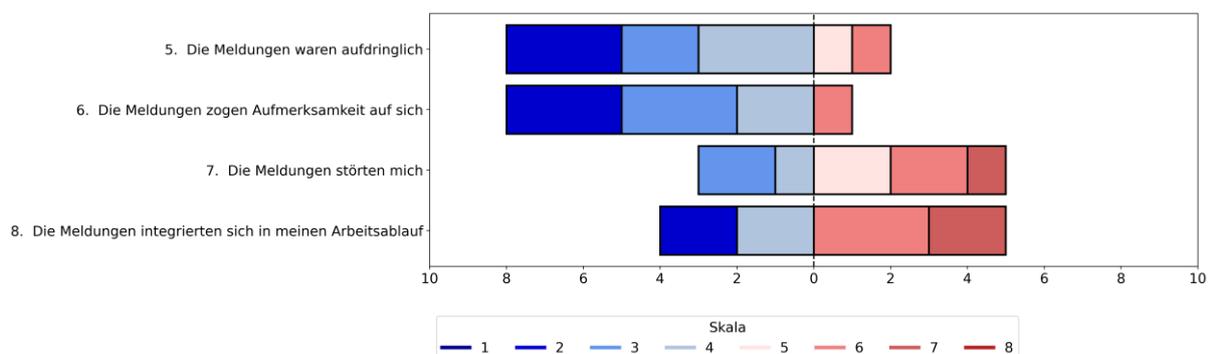


Abbildung 4-56 - Bewertung in Bezug auf den Arbeitsfluss

Abbildung 4-57 zeigt, dass die Mehrheit der Probanden die Meldungen beachteten und sich darauf verlassen haben. Außerdem bewerteten über 60% der Probanden die Interventionen als nützlich.

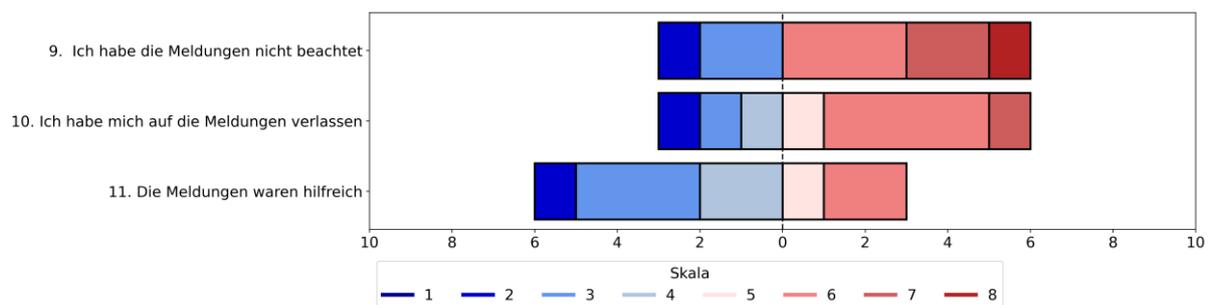


Abbildung 4-57 - Bewertung in Bezug auf Vertrauen und Nützlichkeit

Im Vergleich zum Versuch ohne Interventionen waren 8 Probanden überzeugt, dass sich ihre Leistung verbessert hat. Etwa die Hälfte der Probanden befand, dass die Arbeit mit dem Assistenzsystem frustrierender und belastender war.

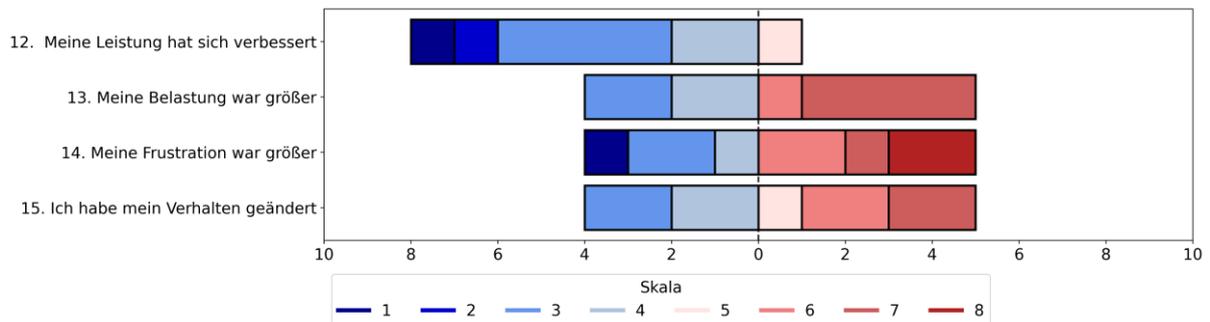


Abbildung 4-58 - Bewertung im Vergleich zum Versuch ohne Interventionen

Die Zuweisung der Attribute zeichnet trotz der eher gemischten Bewertung ein positives Bild. Die Interventionen waren klar verständlich und vorhersehbar. Sie wurden außerdem als eher nachvollziehbar und zuverlässig bewertet. Das System wurde aber auch überwiegend als vorschnell und dumm bezeichnet.

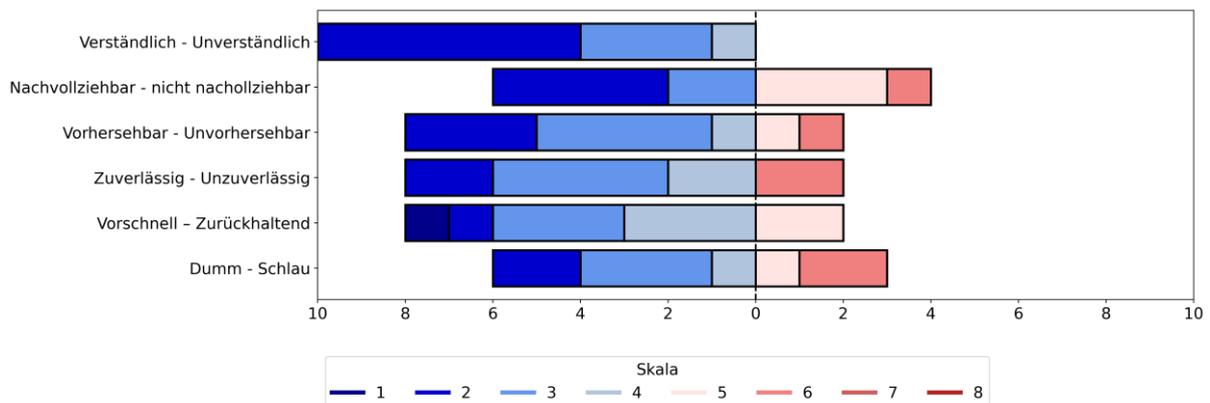


Abbildung 4-59 - Zuweisung von Attributen an die Interventionsgenerierung

In den Gesprächen mit den Probanden wurden die folgenden Aspekte noch hervorgehoben:

- Die Interventionen halfen den Probanden vor allem in zwei Situationen: wenn das auftauchende Objekte außerhalb des aktuellen Kartenbereichs erschien oder bei der Identifikation noch nicht bearbeiteter Gebäude, welches durch die visuell überladene taktische Karte schwerfiel.
- Für die Probanden hatten die Interventionen auf der taktischen Karte einen wesentlich weniger disruptiven Charakter verglichen mit den auditiven Benachrichtigungen.

#### 4.4.3 DISKUSSION

Zusammenfassend lässt sich sagen, dass die SB-Abschätzung aussagekräftige Aspekte der individuellen Leistung des Teilnehmers erfasst und die Ergebnisse in der Kontrollbedingung zeigen, dass es mit Pilotenfehlern im Versuch korreliert ist. In der Bedingung mit adaptiver Intervention konnte die Abschätzung verwendet werden, um sinnvolle Warnungen zu generieren, auch ohne direkt zu prüfen, ob ein Fehler vorliegt. Die Warnungen verbesserten die durchschnittliche Leistung und verringerten die Varianz, insbesondere bei unvorhersehbaren Ereignissen wie beispielsweise der Störung der Flughöhe durch virtuelle Böen. Während die Ergebnisse bei der Verfolgung von Geschwindigkeit und Höhe sowie der Delegationszeit konsistent waren, verschlechterte sich die

Leistung bei der Flugpfadverfolgung. Eine Erklärung hierfür könnte sein, dass die Warnungen für Geschwindigkeit und Höhe die Piloten vom Verfolgen der Route ablenkten. Eine alternative Erklärung ist, dass die Piloten schnell das Vertrauen in die Routenwarnung verloren, da sie manchmal ohne Fehler ausgelöst wurde (vergleiche Abbildung 4-43). Dies hatte zur Folge, dass die Teilnehmer die Flugpfadwarnungen ignorierten. Daraus können zwei Schlussfolgerungen gezogen werden: Erstens sind unvorhersehbare Änderungen von Systemzuständen bessere Ziele für die adaptive Alarmierung als vorhersehbare Änderungen. Daher sollte eine vorhersehbare Änderung von Werten eine größere Verzögerung bis zur Auslösung einer Warnung nach sich ziehen, da eine hohe Wahrscheinlichkeit besteht, dass der Pilot die Änderung kennt. Zweitens sind falsch-positive Meldungen ein kritischer Faktor für das Vertrauen der Benutzer und müssen bei der Entwicklung berücksichtigt werden. Ein ähnliches Problem wurde auch von Lounis, Peysakhovich und Causse (2020a) beobachtet.

Obwohl die Leistung fast aller Teilnehmer gesteigert werden konnte, war die subjektive Bewertung nicht durchweg positiv, wofür zwei Dinge verantwortlich sein könnten. Erstens trugen fehlerhafte Interventionen zu einer negativen Wahrnehmung bei, insbesondere bei der Flugpfadverfolgungswarnung, die oft unnötig ausgelöst wurde. Zweitens wurden die Piloten vor dem Versuch nicht über das Grundprinzip des adaptiven Systems informiert. Dies führte zu Irritationen, wenn die Piloten nicht verstanden, warum es eine Warnung gab. Die Kombination aus einer hohen Anzahl von Fehlalarmen und geringer Transparenz könnte für die differente Bewertung verantwortlich sein. Dies stände auch im Einklang mit der Arbeit von Dorneich et al. (2016), bei der die geringe Transparenz und Vorhersagbarkeit adaptiver Systeme im Cockpit von Luftfahrtfachleuten als größtes Risiko für die Flugsicherheit und die Akzeptanz bewertet wurde. Um die Akzeptanz zu erhöhen, sollten die Teilnehmer daher im Umgang mit dem adaptiven System geschult werden.

Außerdem war die Gestaltung der Benachrichtigungen für den ersten Versuch sehr aufdringlich. Die lautstarken Benachrichtigungen unterbrachen die Teilnehmer in ihrem Arbeitsablauf. Diese Gestaltung könnte sinnvoll sein, um die Aufmerksamkeit auf sicherheitskritische Informationen zu lenken, sollte aber bei weniger relevanten Informationen nicht verwendet werden. Alternativ hätte das System auch im ersten Versuch Adaptionen auslösen können, die nur die Auffälligkeit geänderter Parameter erhöhen (vergleichbar mit Fortmann und Mengerlinghausen (2014)). Ein anderer Ansatz wäre es, die Warnungen zu verzögern, bis der Teilnehmer seine aktuelle Aufgabe beendet hat. Dies setzt wiederum voraus, dass das Ende einer Tätigkeit erkannt werden muss.

### Limitierungen

Die Aussagekraft der Studie hat einige Einschränkungen. Erstens konnte der Trainingseffekt nicht eliminiert werden, da die Versuchsreihenfolge "keine Alarmierung" gefolgt von "adaptiver Alarmierung" lautete. Dies könnte für die Leistungsverbesserungen verantwortlich sein. Dann wäre jedoch fraglich, warum sich die Probanden in der Aufgabe der Flugpfadverfolgung verschlechtern.

Zweitens gab es in der Kontrollbedingung kein Unterstützungssystem. Dies ist ein schlechter Vergleich, da jede Art von Unterstützung eine gute Chance hat, die Leistung im Vergleich zu keiner Unterstützung zu verbessern. Zweifelsohne würde ein System, das im Fehlerfall Alarm schlägt noch wesentlich bessere Ergebnisse erzielen als die SB-adaptiven Benachrichtigungen. Man muss jedoch anerkennen, dass das System ohne eine Kenntnis der genauen Ziele in vielen Fällen sinnhafte Benachrichtigungen generiert hat.

Zuletzt war, wie bereits im ersten Validierungsversuch diskutiert, die Erfahrung der Teilnehmer sehr unterschiedlich, und die Teilnehmer wurden nicht mit dem adaptiven System geschult, was die Leistungsergebnisse verfälschen könnte.

## Zusammenfassung

Aus den Ergebnissen können folgende Aspekte zusammengefasst werden:

- Die SB-Abschätzung kann fehlendes SB robust erkennen, wenn dies unerwartete Informationen betrifft, also Änderungen von Systemwerten aufgrund externer Störungen oder neuartige Informationen mit großen Freiheitsgraden (bspw. Position einer Bedrohung).
- Die Interventionen für diese Fälle können die Leistung des Bedieners steigern, beinhalten aber das Risiko, dass der Bediener von anderen Aufgaben abgelenkt wird (siehe Leistung in der Flugpfadaufgabe).
- Die adaptive Strategie, dem Bediener unbewusste Informationen mitzuteilen, wird dann besonders kritisiert, wenn die Information keine Handlung erfordert.
- Die Akzeptanz des adaptiven Systems leidet unter zu vielen falsch-positive Meldungen.

## 5 VALIDIERUNG DES GESAMTSYSTEMS IN EINER MILITÄRISCHEN LUFTOPERATION

---

Die bisher beschriebenen Ergebnisse in den Validierungsstudien zeigen, dass ein SB-adaptives Auslösen von Benachrichtigungen technisch umsetzbar ist und Fehler des Bedieners in einigen Situationen vermieden werden können. Die für den Abschnitt 4.4 umgesetzten Interventionen waren speziell angepasst auf die Versuchsaufgaben, bei denen die aufgabenrelevanten Informationen direkt im SB-Modell abgebildet werden konnten. Die vorliegende Arbeit wurde im Rahmen eines größeren Forschungsprojekts im Kontext des FCAS-Entwicklungsprogramms für ein neues europäisches Kampfflugzeug entwickelt. Ein Thema war die Gestaltung des zukünftigen Kampfflugzeugcockpits und den Einsatz adaptiver Systeme zur Unterstützung des Piloten, weshalb die Anwendungsdomäne für das folgende Kapitel eine militärische Luftoperation mit Kampfflugzeugen in bemannt-unbemannten Teams ist. Bei der Übertragung des SB-adaptiven Konzepts auf diese Arbeitsdomäne stellt sich die Frage, welche Art von Interventionen für den Piloten eines Kampfflugzeugs in einem solchen Missionskontext nützlich sind.

Um diese Frage zu adressieren, beschreibt das folgende Kapitel eine Methode zur Entwicklung von SB-adaptiver Assistenz in komplexen Arbeitsdomänen. Hierfür wird in Abschnitt 5.1 zunächst die Arbeitsdomäne einer militärischen Luftoperationen mit bemannt-unbemannten Teams eingeführt. Im darauffolgenden Kapitel werden dann die Durchführung und Ergebnisse einer Aufgabenanalyse zur Identifikation von Anwendungsfällen für SB-adaptive Assistenz beschrieben. Abschließend werden die Ergebnisse der Versuche zur Validierung der implementierten Interventionen mit einer kleinen Zahl von Kampfflugzeugpiloten beschrieben.

### 5.1 DIE ARBEITSDOMÄNE: MUM-T IN ZUKÜNFTIGEN KAMPFFLUGZEUGMISSIONEN

*Manned-Unmanned-Teaming* (MUM-T) hat als neue operationelle Einsatzform Einzug in den großen Entwicklungsprogramme zukünftiger luftgestützter Waffensysteme gehalten (Lindner, 2022). MUM-T steht dabei für die

„[...] INTEROPERABILITÄT VON BEMANNTEN UND UNBEMANNTEN PLATTFORMEN[...] ZUR VERFOLGUNG EINES GEMEINSAMEN MISSIONSZIELS INNERHALB DESSELBEN RÄUMLICH, ZEITLICH UND INTENTIONAL BEGRENZTEN KONTEXTES EINER MISSION.“ (BUNDESWEHR INTRANET, 2020)

Das Entwicklungsprojekt Future Combat Air System (FCAS) sieht ein luftgestütztes Waffensystem bestehend aus bemannten und unbemannten Kampfflugzeugen vor. Darin enthalten ist das Konzept, dass die Kampfflugzeugpiloten neben ihren eigenen Waffensystemen, mehrere unbemannte Plattformen aus dem Cockpit heraus führen (Airbus, 2021). Dabei soll das Gesamtsystem vom Einsatz unbemannter Plattformen profitieren, ohne dabei die menschliche Stärke und Flexibilität bei der Entscheidungsfindung und Durchführung einer Mission zu verlieren sowie weiterhin ein hohes Maß an Verantwortbarkeit zu erhalten.

Das Zusammenspiel zwischen bemannten und unbemannten Plattformen stellt hohe Anforderungen an den Bediener und die Konzeption des Gesamtsystems (Schulte et al., 2016). Die technischen Lösungsansätze umfassen neben dedizierten Datenlinks, die Mensch-Maschine-Schnittstelle, Missionsmanagement- und -ausführungsfunktionen, sowie intelligente Unterstützungsfunktionen. Neben diesen Herausforderungen beim Systementwurf steigen auch die Anforderung an den Piloten

(Lindner et al., 2022). Die Führung mehrerer unbemannter Systeme mit hohem Automationsgrad ist eine anspruchsvolle Aufgabe mit stark schwankender Arbeitsbelastung (Brand & Schulte, 2021). Eine kognitionsergonomische Gestaltung des Pilotenarbeitsplatzes, die diese Problematiken adressiert, ist Voraussetzung für ein effektives Gesamtsystem. Für die vorliegende Arbeit wurde auf einem System aufgebaut, das im Rahmen von Forschungsarbeiten am Institut für Flugsysteme eine Umsetzung von MUM-T im Rahmen einer Kampfflugzeugmission evaluieren sollte. Eine detaillierte Beschreibung des Gesamtsystems kann in der Arbeit von Lindner (2022) nachgelesen werden.

Im verwendeten System erhält der Pilot einen Missionsauftrag aus einer übergeordneten *Command-and-Control*-Entität (C2). Eingebettet in der räumlichen, zeitlichen und kontextuellen Umgebung der Mission soll das Missionsziel mit einem Team bestehend aus Kampfflugzeug und unbemannten Plattformen das Missionsziel erreichen, welches beispielsweise eine Bekämpfung oder Aufklärungsleistung sein kann. Der Pilot ist an Bord des Kampfflugzeugcockpits und steuert die eigene Plattform. Parallel hierzu führt er bis zu drei unbemannten Plattformen (*engl. Unmanned Combat Aerial Vehicle, UCAV*) mit. Der Pilot delegiert Aufträge an diese UCAVs mithilfe von dedizierten Missionsplanungsfunktionen. An Bord der UCAVs sind software-basierte Agenten, welche die zugewiesenen Aufträge ausführen wie beispielsweise eine Aufklärung eines Gebäudes. Aus dieser Zielvision resultieren folgende drei Hauptaufgabentypen für einen Kampfflugzeugpiloten:

- |   |
|---|
| <ul style="list-style-type: none"><li>▪ Steuern der eigenen Plattform zur Ausführung einer Missionsaufgabe</li><li>▪ Überwachung der taktischen Lage und des Missionsfortschritts</li><li>▪ Delegation &amp; Anpassung von Aufträgen an die Plattformen des Teams</li></ul> |
|---|

Das SB-adaptive Assistenzsystem soll den Piloten bei diesen Aufgaben unterstützen. Hierfür muss zunächst geklärt werden, welche Informationen für den Piloten in der beschriebenen Arbeitssituation relevant zur Bildung von SB sind. Hierfür wurde eine Aufgabenanalyse durchgeführt, welche im nächsten Abschnitt vorgestellt wird.

## 5.2 GESTALTUNG VON SB-ADAPTIVER ASSISTENZ IN MUM-T

Zur Gestaltung von SB-adaptiver Assistenz braucht es Wissen darüber, welche dynamischen Systeminformationen in einer gegebenen Aufgabensituation relevant sind. Diese Frage ist für eine militärische Luftoperationen nicht trivial und kann nur von Expertenbedienern beantwortet werden. Zur Identifikation der relevanten Aufgabeninformationen, werden im Folgenden Durchführung und Ergebnisse einer zielgetriebenen Aufgabenanalyse mit Kampfflugzeugpiloten beschrieben.

### 5.2.1 AUFGABENANALYSE

Eine Aufgabenanalyse (*engl. Task Analysis*) bezeichnet den Prozess zur Identifikation und Untersuchung von Aufgaben eines Menschen – meist bei der Bedienung eines technischen Systems (Kirwan & Ainsworth, 1992). Zu etablierten Methoden gehört beispielsweise die hierarchische Aufgabenanalyse, deren Ziel es ist, die Arbeit eines Bedieners im Sinne von übergeordneten Zielen, Aufgaben und Unteraufgaben zu modellieren (Stanton, 2006). Aufbauend auf der hierarchischen Aufgabenanalyse wurde die Klasse der kognitiven Aufgabenanalysen entwickelt, um in der Lage zu sein, Wissen und Denkprozesse eines Bedieners beschreiben zu können. Zentrale Motivation hierfür ist, dass die hierarchische Analyse zwar beobachtbare Aspekte der Aufgabendurchführung abbildet, aber nicht die Entscheidungsprozesse des Bedieners integriert (Richard E. Clark et al., 2001).

Als eine Form der kognitiven Aufgabenanalyse für SB, haben Endsley und Jones (2012) die *Goal Directed Task Analysis (GDTA)* entwickelt, welche darauf abzielt, eine Aufgabenumgebung nach den Zielen des Bediener, den Entscheidungen zur Erreichung dieser Ziele und den Informationsanforderungen für diese Entscheidungen zu strukturieren. Die GDTA fokussiert sich dabei auf dynamische Informationen, welche sich während der Arbeitssituation ändern können und welche der Bediener kontinuierlich überprüfen muss. Ferner ist die GDTA eine Analysemethode, die Aufgaben unabhängig von einer technischen Umsetzung beschreibt. Die Struktur einer GDTA ist in Abbildung 5-1 abgebildet und zeigt, dass übergeordnete Ziele und Subziele definiert werden, welche im Zusammenhang mit Entscheidungen des Bediener stehen. Mit den Entscheidungen werden wiederum Information assoziiert, die für die Entscheidungsfindung relevant sind. Diese Art der Aufgabenanalyse wurde unter anderem in Domänen der zivilen Luftfahrt (Endsley & Jones, 2012), Heeresführung (Bolstad et al., 2002), Schifffahrt (Haffaci et al., 2021) oder Notfallmedizin (Hamid & Waterson, 2011) durchgeführt.

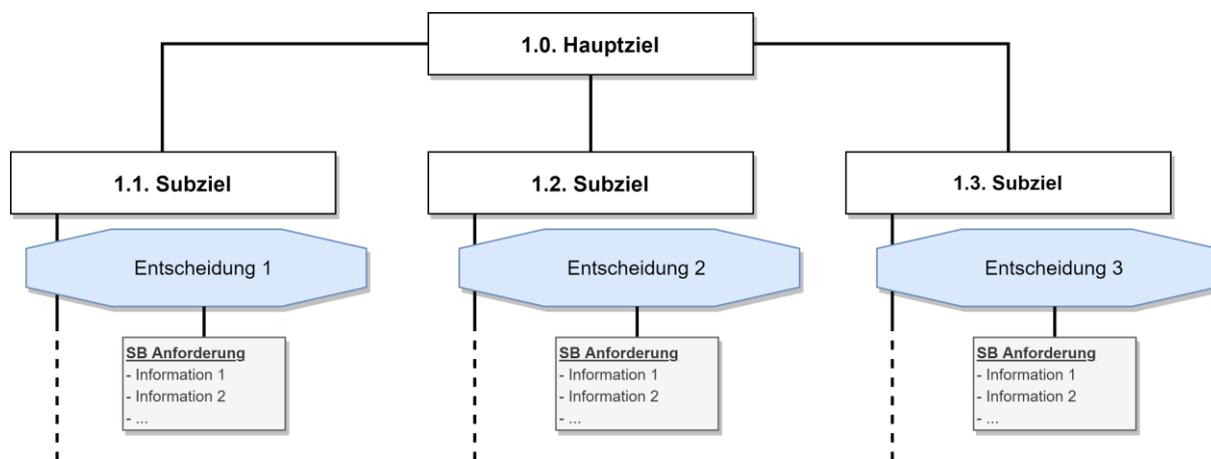


Abbildung 5-1 - Generische Struktur einer GDTA (Endsley & Jones, 2012)

Wesentlicher Bestandteil einer GDTA ist die Durchführung von Interviews mit Fachexperten, um mit deren Hilfe alle Bestandteile der Aufgabenstruktur zu identifizieren. Die Methode basiert auf unstrukturierten Befragungen, welche durch Analyse von schriftlichen Arbeiten und Dokumentation aus der Domäne ergänzt werden. Die Autoren haben die GDTA entwickelt, um es Gestaltern von MMS einfacher zu machen, die Aufgaben eines Bediener und die dafür notwendigen Informationen zu verstehen. Aus diesem Grund ist sie auch für die vorliegende Fragestellung nach den relevanten Informationen für den Piloten eines Kampfflugzeugs besonders geeignet, da sie die zentrale Fragestellung nach relevanten Informationen direkt beantwortet. Für die vorliegende Arbeit wurde eine GDTA mit Kampfflugzeugpiloten durchgeführt, deren Ergebnisse im Folgenden beschrieben wird.

#### 5.2.1.1 Versuchsteilnehmer & Ablauf

Die Piloten wurden im Rahmen der in Abschnitt 4.3 durchgeführten Studie befragt, wobei hier nur die acht professionellen Kampfflugzeugpiloten der deutschen Luftwaffe herangezogen wurden (siehe Tabelle 4-7). Unter den Probanden waren Piloten des Waffensystems Tornado, welches vorwiegend für die Einsatzformen *Air Interdiction (A/I)* und *Suppression of Enemy Air Defence (SEAD)* eingesetzt wird, sowie Piloten des Waffensystem Eurofighter, welches für Luftkampf vorgesehen ist. Die Befragung fand im Rahmen einer einwöchigen Versuchskampagne im Forschungssimulator am Institut für Flugsysteme der Universität der Bundeswehr statt. Für das Interview wurden die Piloten im Vorfeld über die Methodik und das Ziel der Befragung aufgeklärt. Es wurde außerdem ein Missionsbriefing einer *Air Interdiction* Mission vorbereitet, um den Einstieg in die Befragung zu erleichtern und das

Gespräch anhand von Missionsphasen zu strukturieren. Auf der Grundlage des Briefings wurden die verschiedenen Missionsphasen durchgesprochen und dabei versucht, relevante Ziele, Subziele und damit verbundene Entscheidungen zu identifizieren. Sobald eine Entscheidung identifiziert wurde, sollte der Proband alle relevanten Informationen benennen, die mit dieser Information verbunden sind. Die Befragung jedes Piloten dauerte etwa zwei Stunden. Nach den Interviews wurden Ziele, Entscheidungen und Informationen in einer Baumstruktur eingeordnet. Es ist anzumerken, dass die befragten Piloten nur Aussagen über die Delegation von Aufträgen an andere menschliche Piloten aus ihrem bemannten Team machen konnten, da bisher kein bemannt-unbemanntes Team realisiert ist.

## 5.2.2 ERGEBNISSE

Ein Überblick über die erzielte Aufgabenstruktur ist in Abbildung 5-2 dargestellt. Das vollständige Ergebnis der Analyse ist im Anhang E.5 dokumentiert. Insgesamt wurden 46 Ziele beziehungsweise Subziele und 63 Entscheidungen identifiziert.

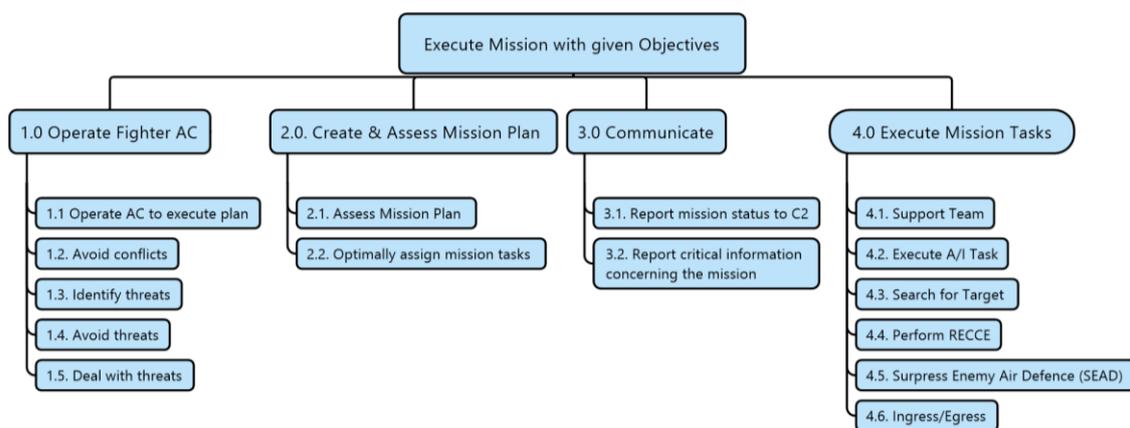


Abbildung 5-2 - Übergeordnete Struktur der Ziele und Sub-Ziele der GDTA

Unter allen Zielen wurden zunächst vier wesentliche Subziele festgestellt. Das erste Subziel „*Operate Fighter AC*“ bezieht sich auf die Führung des eigenen Kampfflugzeugs im Missionsgebiet mit dem Ziel eines sicheren Flugpfads sowie der Identifikation von und Reaktion auf Bedrohungen für die eigene Plattform. Das zweite Sub-Ziel „*Create & Assess Mission Plan*“ beschreibt Ziele im Zusammenhang mit der Verteilung von Aufgaben an Teammitglieder. Beim dritten Subziel „*Communicate*“ geht es um die Kommunikation mit C2 und dem eigenen Team, wofür kontinuierlich evaluiert werden muss, ob der aktuelle Ablauf der Missionsplanung entspricht oder es eine missionskritische Änderung gab, die kommuniziert werden muss. Das vierte Subziel „*Execute Mission Task*“ fasst alle Ziele zusammen, die mit der aktiven Ausführung einer Missionaufgabe aus dem Spektrum *Supression of Enemy Air Defense* (SEAD) oder *Air Interdiction* (A/I) resultieren. Die Aufgaben Luftkampf oder elektronischer Kampf wurden in dieser Analyse nicht speziell abgefragt.

Die Entscheidungen und Informationsanforderungen befinden sich an unterster Ebene der Baumstruktur. Abbildung 5-3 zeigt hier ein Beispiel, wie das Sub-Ziel „*Identify Threats*“ heruntergebrochen und die Evaluierung wahrscheinlicher sowie existierender boden- und luftgestützter Bedrohungen unterteilt. Das Unterziel „*Asses ground threats*“ (1.3.2) zeigt, welche Informationen notwendig sind, um die Bedrohung einer bekannten bodengestützte Luftverteidigung zu bewerten.

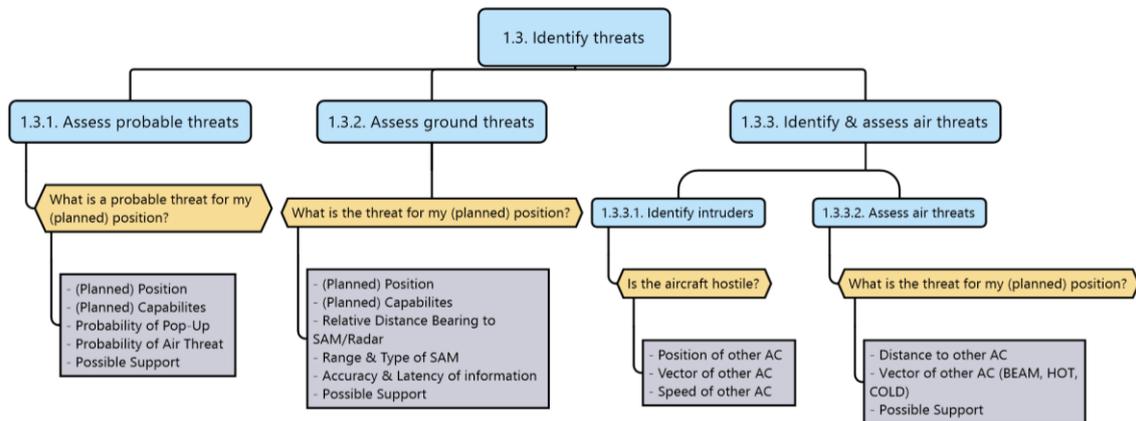


Abbildung 5-3 - Sub-Ziel "Identify Threats"

Mit den Ergebnissen der GDTA können viele der relevanten, dynamischen Informationen für einen Kampfflugzeugpiloten in einer A/I-Mission identifiziert werden. Auf Befragung im Zusammenhang mit MUM-T wurde verzichtet, da die Piloten noch keine operationelle Erfahrung mit dieser Einsatzart hatten. Zudem wurde darauf geachtet, Aufgaben und Informationen unabhängig von ihrer Umsetzung in einem bestehenden Waffensystem (bspw. dem Eurofighter) zu beschreiben, um die Weiterverwendung der Information im Forschungskontext zu gewährleisten.

### Limitierungen

Die durchgeführte Aufgabenanalyse hat einige Schwachstellen. Zunächst wurde nur die Missionsaufgaben Luft-Boden-Bekämpfung und Unterdrückung von Luftabwehr betrachtet, um die Befragung durchzuführen. Aspekte der Missionsaufgaben Luftkampf oder elektronischer Kampf wurden vernachlässigt. Außerdem gab es in der Befragung zu mehreren Zeitpunkten die Situation, dass einige Information als geheim eingestuft waren und eine weitere Befragung nicht möglich war. Zudem wurde nur eine Befragungsrunde durchgeführt und das Ergebnis der Aufgabenanalyse nicht nach einer Strukturierung durch Experten validiert. Aus diesem Grund kann kein Anspruch der Vollständigkeit erhoben werden, aber trotz dieser Schwachstellen konnte eine aussagekräftige, initiale Strukturierung des Aufgaben- und Informationsraums erzielt werden. Dieses Wissen wird im Folgenden verwendet, um sinnvolle Anwendungsfälle für die SB-adaptive Assistenz zu finden.

### 5.2.3 AUSWAHL & IMPLEMENTIERUNG VON ANWENDUNGSFÄLLEN

Für die Validierung des Gesamtsystems wurden Entscheidungen aus der Aufgabenanalyse ausgewählt und hierfür SB-adaptive Benachrichtigungen für den Simulator (siehe Abschnitt 4.1) implementiert. Dabei wurde auf den vollen Funktionsumfang des Simulators zurückgegriffen. Die Informationen, die mit diesen Entscheidungen in Verbindungen stehen, wurden dann im SB-Modell umgesetzt und visuelle sowie auditive Interventionen realisiert. Das SB-Modell, das als Wissensbasis für die Interventionen eingesetzt wurde, ist in Abbildung 5-4 dargestellt. Im Folgenden werden die umgesetzten Interventionen im Zusammenhang mit dem zugehörigen Ziel aus der GDTA beschrieben. Ergänzend zu den Anwendungsfällen aus der GDTA wurden Interventionen implementiert, die für das Missionsmanagement im Kontext von MUM-T wichtig sind. Die durchgeführte GDTA konnte hierfür keine Erkenntnisse liefern, wobei einige Entscheidungen von der Führung des eigenen Flugzeugs (Subziel 1.0) oder der Führung von Teammitgliedern (Subziel 3.0) übertragen werden konnten.

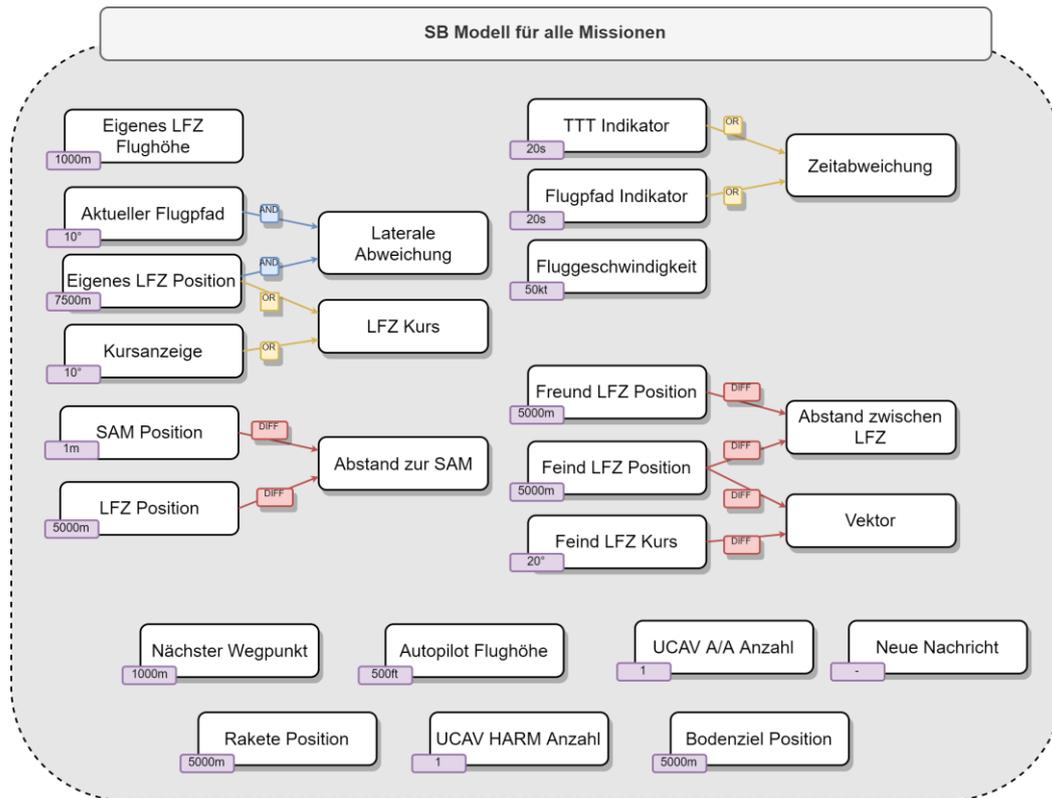


Abbildung 5-4 - SB-Modell als Wissensbasis für die Interventionen. Annotiert an den Knoten sind die Normierungskonstanten dargestellt (lila).

Voraussetzung für die Umsetzung eines Assistenzfalls zur Validierung in der virtuellen Simulation ist, dass die verwendete Systemschnittstelle, in diesem Fall der Forschungssimulator, die Information auch repräsentiert. Wenn die Information nicht repräsentiert ist, gibt es auch keinen Systemwert, der als Vergleich bei der SB-Bewertung herangezogen werden kann.

Basierend auf den Bewertungen der Piloten aus dem vorherigen Versuch zur Interventionsgenerierung wurden der Adaptionmechanismus angepasst. Vor Auslösen einer Intervention bei hoher SB-Abweichung wird geprüft, ob eine kritische Bedingung gegeben ist, die eine Intervention rechtfertigt, wie beispielsweise eine Abweichung von der geplanten Flughöhe. Informationen über Missionsplan und Situation sind durch den Systemaufbau (siehe Abbildung 3-11) in der Interventionsgenerierung verfügbar und wurden anwendungsfallspezifisch verarbeitet. Im Folgenden werden nun die implementierten Interventionen beschrieben und in Bezug zu den Subzielen der GDTA gestellt.

### Maintain desired flight path (GDTA Sub-Ziel 1.1.1.)

Für das Ziel, den geplanten Flugfad zu halten, wurde eine Intervention für die Fluggeschwindigkeit, Flughöhe, laterale Flugfadabweichung sowie die Abweichung von der geplanten Missionszeit implementiert. Alle Interventionen wurden über die Sprachsynthese-Schnittstelle als Sprachbenachrichtigung an den Probanden herangetragen. Tabelle 5-1 zeigt die generierten Sprachnachrichten sowie die Bedingungen für einen kritischen Zustand.

Intervention	Sprachnachricht	Bedingung kritischer Zustand	Time-until-notification [s]
Flughöhe	“Flight Level <X> Target Flight Level <X>”	Abstand zur Zielhöhe > 500 Fuß	8
Fluggeschwindigkeit	„Speed <X>“	Early/Late > 15s	10
Laterale Flugpfadabweichung	„High route deviation“	Lateraler Abstand zur Route > 4000m	10
Zeitabweichung	„Early/Late <X> Seconds“	Early/Late > 15s	5
Neue Nachricht	„New Message: <Message>“	-	5

Tabelle 5-1 - Interventionen für das Ziel „Maintain desired flight path

Die Benachrichtigung für die laterale Abweichung wurde noch ergänzt durch eine saliente Markierung des aktuellen Flugpfads in der taktischen Karte.

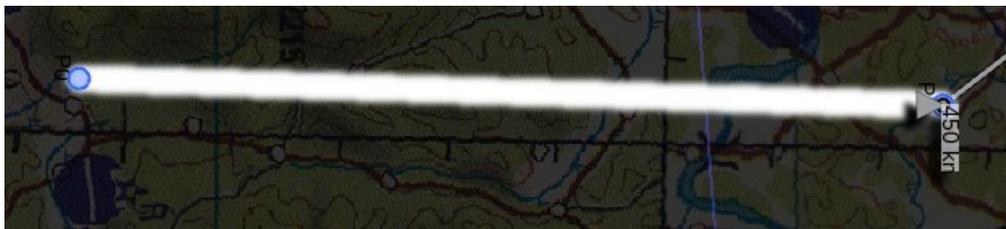


Abbildung 5-5 - Markierung des aktuellen Flugpfads bei Intervention

Die Abbildung 5-4 zeigt das für diese Interventionen verwendete SB-Modell. Die Interventionsgenerierung liest die Abweichungswerte in den Knoten „Flugzeughöhe“, „Laterale Abweichung“, „Zeitabweichung“, „Fluggeschwindigkeit“ und „Neue Nachricht“ für das eigene Flugzeug aus und generiert nach dem in Abschnitt 3.5.2.5 beschriebenen Mechanismus eine Intervention, sofern die Bedingungen für eine kritische Situation vorliegen. Das heißt, dass diese Interventionen nicht für unbemannte Plattformen generiert werden.

### Manage Automation (GDTA Sub-Ziel 1.1.3.)

Die Ergebnisse der GDTA haben im Zusammenhang mit der Flugführungsautomation nur Ergebnisse geliefert, die Anhaltspunkte liefern, wann der Pilot den Autopiloten benutzt. Hier gab es keine Ergebnisse im Hinblick auf das notwendige SB während der (Wegpunkt-) Autopilot aktiviert ist. Studien im Zusammenhang mit dieser Art von Cockpitautomation haben aber gezeigt, dass sich Piloten oft nicht dem aktuellen Modus des Autopiloten bewusst sind oder ihre Überwachungsaufgabe

vernachlässigen (Dehais et al., 2015; Sarter et al., 2007). Aus diesem Grund wurden Interventionen geschaffen, die keine kritische Situation voraussetzen, sondern den Zustand der Automation an den Piloten melden, sofern er die relevanten Zustände nicht überwacht. Tabelle 5-2 zeigt die umgesetzten Interventionen mit den Bedingungen, unter denen sie ausgelöst werden können.

Intervention	Sprachnachricht	Bedingung kritischer Zustand	Time-until-notification [s]
<b>Autopilot Flughöhe</b>	“Autopilot reached target altitude <X>”	Autopilot ALT aktiv & Zielhöhe erreicht	3
<b>Autopilot Fluggeschwindigkeit</b>	“Autopilot reached target velocity <X>”	Autothrottle aktiv & Zielgeschwindigkeit erreicht	3
<b>Autopilot Kurs</b>	“Autopilot reached target heading <X>”	Autopilot HDG aktiv & Zielkurs erreicht	3
<b>FMS Flughöhe</b>	„FMS target altitude set to <X>”	FMS aktiv	3
<b>FMS Wegpunkt</b>	„FMS waypoint reached. New heading set to <X>”	FMS aktiv	3

Tabelle 5-2 - Interventionen für das Ziel „Manage Automation“

Abbildung 5-4 zeigt die zusätzlichen Knoten des SB-Modells zur Umsetzung dieser Funktionalität. Die Autopilot-Interventionen lesen den Wert der „Flughöhe“, „Fluggeschwindigkeit“ und „Flugzeugkurs“ aus. Der Autopilot gibt somit eine Benachrichtigung, wenn er vom Modus *Acquire* in den Modus *Hold* übergeht. Die Interventionen im Zusammenhang mit dem Wegpunktautopiloten (hier *Flight Management System, FMS*, genannt) lesen den Knoten „Autopilot Flughöhe“ und „Nächster Wegpunkt“ aus. Hier ist die Idee den Piloten zu benachrichtigen, wenn das FMS den Autopiloten neu konfiguriert, um den nächsten Wegpunkt beziehungsweise eine neue Flughöhe einzunehmen.

#### Assess ground threats (GDTA Sub-Ziel 1.3.2.)

Für das Sub-Ziel „Assess ground threats“ ist vor allem das Bewusstsein über Position und Typ der SAM-Site relevant. Dies wurde in zwei Interventionen umgesetzt, welche in der Tabelle 5-3 beschrieben sind. Die Interventionen wurden als mithilfe der Benachrichtigungen auf der taktischen Karte umgesetzt (siehe Abbildung 4-40).

Intervention	Benachrichtigung	Bedingung kritischer Zustand	Time-until-notification [s]
<b>Neue bodengestützte Bedrohung</b>	Benachrichtigung SAM-Site auf takt. Karte (siehe Abbildung 4-40) mit Text „Ground Threat“	-	3
<b>Bedrohung nahe eigenem Flugzeug</b>	Benachrichtigung SAM-Site auf takt. Karte (siehe Abbildung 4-40) mit Text „Threat“	Zeit bis Flugzeug innerhalb max. Reichweite < 10s	1

Tabelle 5-3 - Interventionen für das Ziel "Assess ground threats"

Für die erste Intervention „Neue bodengestützte Bedrohung“ gibt es keine Bedingung, da eine Bedrohung stets relevant und kritisch ist. Für diese Intervention wird die Abweichung des SB-Knotens „SAM Position“ gelesen, welche aufgrund der statischen Position nur die Abweichungswerte „Nicht bewusst“ ( $x_{dev}^{SAM} = -1$ ) oder bewusst ( $x_{dev}^{SAM} = 0$ ) einnehmen kann. Anders ist es bei der Intervention „Bedrohung nahe eigenem Flugzeug“, für das der Knoten „Abstand zu SAM“ gelesen wird. Dieser Abweichungswert kann sich kontinuierlich ändern durch die Bewegung des Flugzeugs.

#### Assess air threats (GDTA Sub-Ziel 1.3.3.2.)

Das Sub-Ziel „Assess air threats“ wurde ähnlich zu den bodengestützten Bedrohungen umgesetzt mit der Ergänzung einer Intervention für den Fall, dass sich der Kurs eines gegnerischen Flugzeugs so ändert, dass er auf eines der eigenen Flugzeuge zufliegt, im Jargon der Luftwaffe als „HOT“ bezeichnet. Die Interventionen und Bedingungen sind in Tabelle 5-4 aufgeführt.

Intervention	Benachrichtigung	Bedingung kritischer Zustand	Time-until-notification [s]
<b>Neue luftgestützte Bedrohung</b>	Benachrichtigung SAM-Site auf takt. Karte (siehe Abbildung 4-40) mit Text „Air Threat“	-	3
<b>Bedrohung HOT auf Fighter</b>	Benachrichtigung Foe AC auf takt. Karte (siehe Abbildung 4-40) mit Text „HOT“	Distanz < 40km und rel. Kurs Foe-AC < 45°	3
<b>Bedrohung HOT auf UCAV</b>	Benachrichtigung Foe AC auf takt. Karte (siehe Abbildung 4-40) mit Text „HOT“	Distanz < 40km und rel. Kurs Foe-UCAV < 45°	5

Tabelle 5-4 - Interventionen für das Ziel "Assess air threats"

Für die Generierung „Neue luftgestützte Bedrohung“ wird der Knoten „Feind LFZ Position“ verwendet, für die beiden Interventionen „HOT“ wird der Knoten „Vektor“ für das Kampfflugzeug beziehungsweise für das UCAV verwendet.

#### Optimally assign mission tasks (GDTA Sub-Ziel 2.2)

Von den relevanten Informationen, die im Zusammenhang mit der Missionsplanung identifiziert wurden, sind einige schon durch die Sub-Ziele der Bedrohungsbewertung abgedeckt. Ergänzend zu den Benachrichtigungen bei Bedrohung, wurde eine Intervention für neue Bodenziele eingeführt.

Darüber hinaus ist die Bewaffnung als Beschreibung der Fähigkeiten der Teammitglieder eine zentrale Information, die noch nicht abgedeckt ist. Hierfür wurden zwei spezifische Interventionen erstellt, wie in Tabelle 5-5 aufgeführt.

Intervention	Benachrichtigung	Bedingung kritischer Zustand	Time-until-notification[s]
<b>Neues Bodenziel</b>	Benachrichtigung Missile auf takt. Karte (siehe Abbildung 4-40) mit Text „New Building“	-	10
<b>Waffe gefeuert</b>	Benachrichtigung Missile auf takt. Karte (siehe Abbildung 4-40) mit Text „HARM/AMRAAM fired“	-	3
<b>UCAV niedrige Waffenzahl</b>	Benachrichtigung UCAV auf takt. Karte (siehe Abbildung 4-40) mit Text „Low HARM/AMRAAM: X“)	Anzahl der Waffen <3	6

Tabelle 5-5 - Interventionen für das Ziel "Optimally assign mission tasks "

Für die Bewertung, ob der Bediener ein neues Bodenziel kennt, wurde der Knoten „Bodenziel Position“ ausgewertet. Die zweite Intervention liest den Knoten „Missile Position“ und die dritte Intervention liest die Knoten „UCAV A/A Anzahl“ beziehungsweise „UCAV HARM Anzahl“, welche jeweils die Anzahl der Luft-zu-Luft- beziehungsweise HARM-Bewaffnung beschreibt.

### 5.3 VALIDIERUNG DER SB-ADAPTIVEN ASSISTENZ

Eine Problematik im Zusammenhang mit der Frage der Validierung von Assistenzsystemen wurde von Rouse (1988) beschrieben, der bei der Entwicklung von adaptiven Unterstützungssystemen von der Gefahr spricht, dass die isolierte Evaluierung einer Unterstützungsfunktion meist positiv ist, während

“[...] IN REALISTICALLY COMPLEX SYSTEMS, TASKS USUALLY OCCUR AS COMPONENTS OF CONTEXTUALLY RICH MULTI-TASK SITUATIONS. UNAIDED HUMAN PERFORMANCE IN SUCH SITUATIONS DEPENDS ON THE MIX OF TASK DEMANDS AT ANY POINT IN TIME. AS A RESULT, AIDING FOR ANY PARTICULAR TASK MAY SOMETIMES BE OF VALUE AND AT OTHER TIMES BE UNNECESSARY.” (ROUSE, 1988)

Es besteht also die Möglichkeit, dass eine isolierte Assistenzfunktion in Validierungsversuchen gute Ergebnisse erzielt, in realen Situationen aber durch situative Einflüsse und die Aufgabenlast sogar störend ist. Aus diesem Grund wurde für die im folgenden Versuchsmissionen auf ein möglichst realistisches Zusammenspiel mehrerer Aufgaben geachtet. Die leitenden Forschungsfragen waren dabei folgende:

- Ist die Auswahl der Benachrichtigungen sinnvoll und sind die Informationen relevant in einem realistischen Aufgabenkontext?
- Sind die SB-adaptiven Benachrichtigungen nützlich im Kontext einer komplexen militärischen Mission?

### 5.3.1 GESTALTUNG DER STUDIE

Zur Beantwortung der Forschungsfragen wurden qualitative Methoden eingesetzt, die primär auf der subjektiven Bewertung durch Fachexperten und der Beobachtung der Systemnutzung basieren. Das SB-adaptive System wurde in Versuchen mit drei Kampfflugzeugpiloten der deutschen Luftwaffe evaluiert.

#### 5.3.1.1 Versuchsgestaltung

Im Folgenden werden die Missionen, Messgrößen, Versuchsteilnehmer und der Ablauf der Mission beschrieben. Die Studie bestand aus drei Versuchen. Die ersten beiden Versuche zielten auf Evaluierung der einzelnen Anwendungsfälle ab, während in der dritten Mission das Zusammenspiel aller Anwendungen relevant war.

#### Mission 1: Aufklärung

In der ersten Mission führten die Probanden nur das eigene Flugzeug. Dabei sollten sie einem Flugpfad folgen (siehe Abbildung 5-6), während sie über die Textnachrichtenanzeige Anweisungen über die zu fliegende Höhe oder zu verwendende Automation bekamen. Die Nachrichten sind in der Appendix Tabelle 7 im Anhang aufgeführt. Im zweiten Teil des Flugs tauchten Gebäude auf, die sie mit dem optischen Aufklärungssensor des Kampfflugzeugs betrachten und dabei Fotos aufnehmen sollten. Relevante Interventionstypen in dieser Mission waren die im Zusammenhang mit dem Sub-Ziel 1.1.1. und 1.1.3. stehenden Anwendungsfälle sowie die Intervention „Neues Bodenziel“. Der Versuch dauerte etwa 15 Minuten.

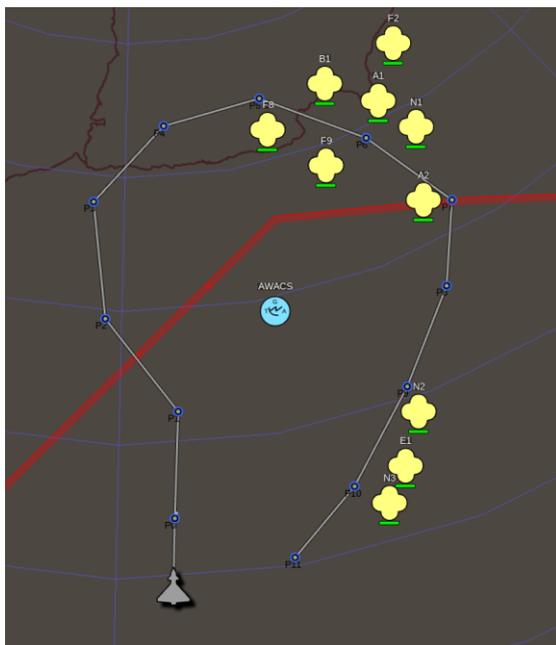


Abbildung 5-6 – Übersicht über die Mission 1



Abbildung 5-7 – Übersicht über die Mission 2

#### Mission 2: Multi-UAV Führung und Überwachung des Luftraums

In der zweiten Mission, dargestellt in Abbildung 5-7, führte der Proband drei UAVs ohne selbst zu fliegen. Dabei sollten 8 Aufklärungsflächen (gelb) abgesucht und gefundene Objekte vom Bediener klassifiziert werden. SAM-Sites, deren Reichweite über die aufzuklärenden Flächen reichten, mussten bekämpft werden. Ebenso sollte der Luftraum überwacht und gegnerische Flugzeuge, die über die gelbe Linie hinaus in das Missionsgebiet fliegen, bekämpft werden. Die gegnerischen Flugzeuge

erschieden in einem Abstand von 1-2 Minuten und hielten sich einige Zeit außerhalb des Gebiets auf bis sie die gelbe Linie überflogen. Die exakten Missionszeiten sind in der Appendix Tabelle 8 dokumentiert. Die Interventionstypen fokussierten sich hier auf die im Zusammenhang mit den Subzielen 1.1.3, 1.3.2, 1.3.3.2 und 2.2 beschriebenen Entscheidungen. Diese Mission dauerte etwa 22 Minuten.

### Mission 3: MUM-T Air Interdiction

Die dritte Mission war angelehnt an eine A/I-Mission mit dem Ziel der zeitgleichen Bekämpfung der Gebäude T1, T3 und T4. Hierfür mussten zunächst die Flugabwehrstellungen bekämpft und eine Aufklärung der Ziele durchgeführt werden. Nach der Klassifikation musste dann eine Bekämpfung der Ziele delegiert beziehungsweise selbst durchgeführt werden. Die Durchführung der Mission wurde bedroht durch drei gegnerische Kampfflugzeuge, die im Verlauf der Mission erschienen, sowie zwei vorher nicht bekannte Flugabwehrstellungen. Die Mission 3 dauerte insgesamt etwa 25 Minuten.

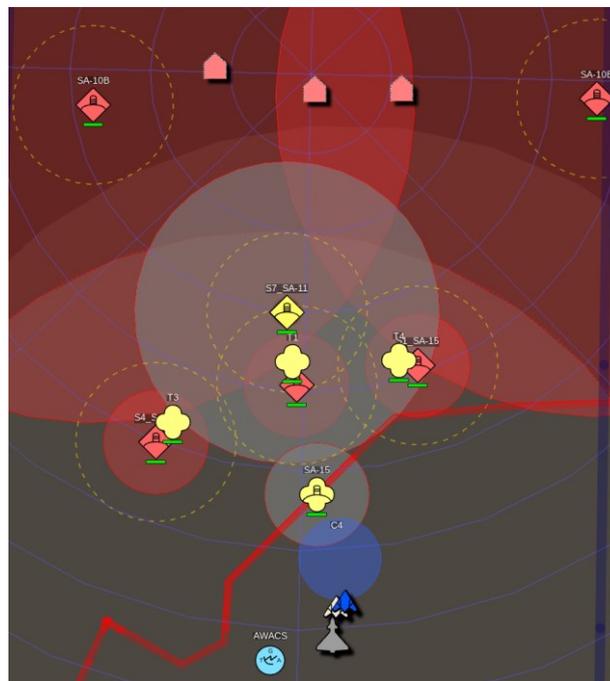


Abbildung 5-8 – Übersicht über die Mission 3.

#### 5.3.1.2 Messgrößen

Zentral für diesen Versuch ist die subjektive Bewertung der Intervention durch die Piloten mittels Fragebögen. Die Versuchsgestaltung sah keinen Vergleich mit einem nicht-adaptiven System vor, weshalb auf eine quantitative Auswertung von Missionsleistung oder Pilotenverhalten verzichtet wurde.

Zur subjektiven Bewertung wurde ein Fragebogen erstellt, mit dem jeder umgesetzte Anwendungsfall einzeln evaluiert wurde. Für die Intervention „Flughöhe“ ist der Fragebogen im Anhang E.6.3 dokumentiert. Der Fragebogen beinhaltet sechs Fragen, die jede Intervention auf die Kriterien Nützlichkeit, Zeitpunkt, Frequenz, Salienz, Richtigkeit sowie Zuverlässigkeit bewerten sollten. Zusätzlich gab es noch offene Fragen zur Bewertung und Verbesserungsmöglichkeiten der spezifischen Assistenz. Außerdem wurde eine allgemeiner Bewertungsbogen erstellt, der eine Gesamtbewertung des Assistenzsystems vorsah. Hierbei war wichtig, inwiefern das Zusammenspiel aus mehreren Interventionen funktionierte und ob das System im Gesamten als nützlich empfunden wurde. Grund

hierfür war wie bereits erwähnt, dass eine positive Evaluierung eines isolierten Anwendungsfalls nicht automatisch eine positive Evaluierung des Zusammenspiels mit anderen Anwendungsfällen bedeutet.

Vier weitere Fragen im Fragebogen zielten auf die Bewertung der Mission mit Hinblick auf den Kräfteansatz, Repräsentativität des Szenarios und der Aufgaben sowie Arbeitsbelastung ab. Dies sollte überprüfen, ob die gewählten Versuchsszenarien eine realistische Mission und Aufgabensituation darstellten.

### 5.3.1.3 Teilnehmer

Teilnehmer des Versuchs waren drei ausgebildete Kampfflugzeugpiloten. Wie Tabelle 5-6 aufführt, sind die Teilnehmer zwischen 35 und 37 Jahren und haben zwischen 370 und 711 Flugstunden auf einem Kampfflugzeug. Alle Teilnehmer waren männlich.

Teilnehmer ID	Alter	Flugstunden	Flugstunden Kampfflugzeug
P1	35	1797	370
P2	37	2000	711
P3	36	1000	700

Tabelle 5-6 - Alter und Flugstunden der Teilnehmer

### 5.3.1.4 Ablauf

Der Versuch wurde bei einer zweitägigen Veranstaltung mit Kampfflugzeugpiloten im Simulator am IFS durchgeführt. Im Rahmen dieser Veranstaltung erhielten die Piloten mindestens zwei Stunden Training im Simulator, bevor sie den hier beschriebenen Versuch durchführten. Für die Versuche mit dem Assistenzsystem gab es kein weiteres spezifisches Training. Vor der Durchführung des Versuchs wurde die Funktionsweise des Assistenzsystems ausführlich erläutert und es wurden alle Anwendungsfälle beschrieben. Im Anschluss an die Erklärung gab es vor jedem Versuch ein Missionsbriefing, in dem Ziele und Randbedingungen der Mission erklärt wurden. Im Anschluss an die ersten beiden Versuche wurden von den Probanden jeweils die anwendungsfallspezifischen Fragebögen sowie der allgemeine Fragebogen ausgefüllt. Nach der dritten Mission wurde lediglich der allgemeine Fragebogen ausgefüllt.

## 5.3.2 ERGEBNISSE

Die Ergebnisse des Versuchs bestehen einerseits aus der Bewertung des Systems in den Fragebögen in 5.3.2.1 und andererseits aus der Beschreibung von Situationen in 5.3.2.2, die das Verhalten des adaptiven Systems im Zusammenhang mit der Arbeitssituation phänomenologisch beschreiben.

### 5.3.2.1 Fragebögen

Im Folgenden werden die Ergebnisse des Fragebogens vorgestellt. Dabei lag die Skala der Antworten stets zwischen 1 („trifft überhaupt nicht zu“) und 5 („trifft voll und ganz zu“).

#### Versuchsbewertung

Der Kräfteansatz, welcher sich auf die Anzahl der Plattformen und deren Fähigkeiten mit Blick auf die Mission bezog, wurde moderat bis gut bewertet. Die Versuchsszenarien und das Aufgabenspektrum wurde nur teilweise als repräsentativ bezeichnet, da der Pilot 3 seine Antwort auf die Führung von UCAVs bezog, welche noch nicht Teil des Aufgabenspektrums eines Kampfflugzeugpiloten ist. Die anderen beiden Piloten bewerteten vor allem bewerteten Mission 2 und 3 als eher repräsentativ.

Mission 1 und 2 wurden von allen Piloten als gut zu bewältigen beschrieben, während Mission 3 schwieriger für Pilot 2 und 3 waren.

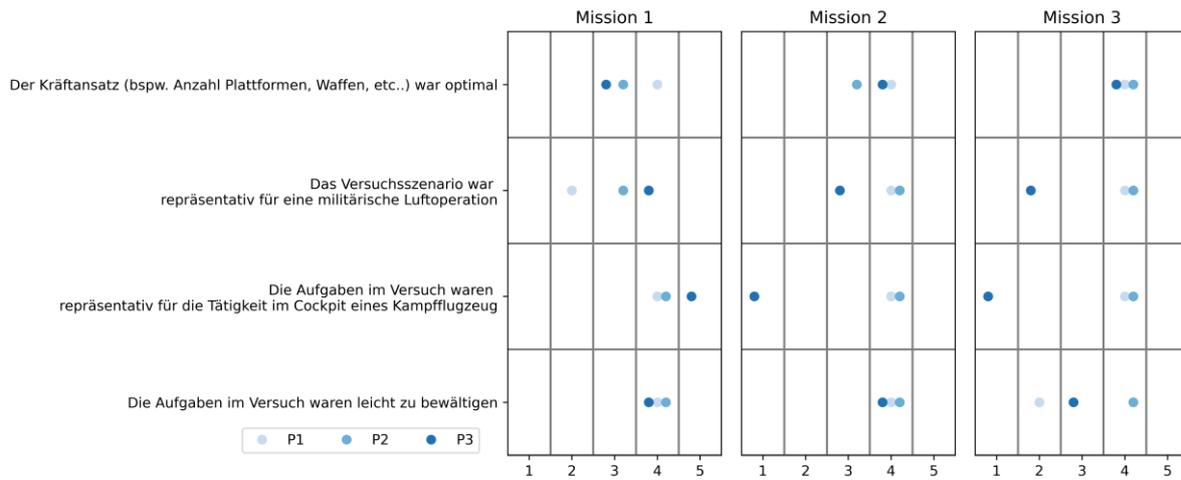


Abbildung 5-9 - Bewertung des Realismus der Missionen (1: „Trifft überhaupt nicht zu“, 5:„Trifft voll und ganz zu“)

Der erste Proband äußerte Kritik bezüglich der Repräsentativität des Aufgabenspektrums, da keinerlei Funkverkehr simuliert wurde. Nach Aussage des Probanden ist dies ein relevanter Aufgabentyp, der mentale Ressourcen bindet und auf dem gleichen auditiven Wahrnehmungskanal liegt wie einige der Benachrichtigungen.

### Interventionsbewertung

Im Folgenden werden die Ergebnisse der einzelnen Anwendungsfälle beschrieben. Wenn auf der Bewertungsskala ein Messpunkt fehlt, bedeutet das, dass der Versuchsteilnehmer sich entweder nicht an die Intervention erinnern konnte oder diese gar nicht auftrat.

Die Abbildung 5-10 zeigt die Bewertung der Interventionen in vier Anwendungsfällen im Bereich des Ziels „Maintain desired flight path (1.1.1.)“, welche nach der ersten Mission abgegeben wurden. Sowohl Flughöhe als auch Fluggeschwindigkeit zeigten heterogene Ergebnisse.

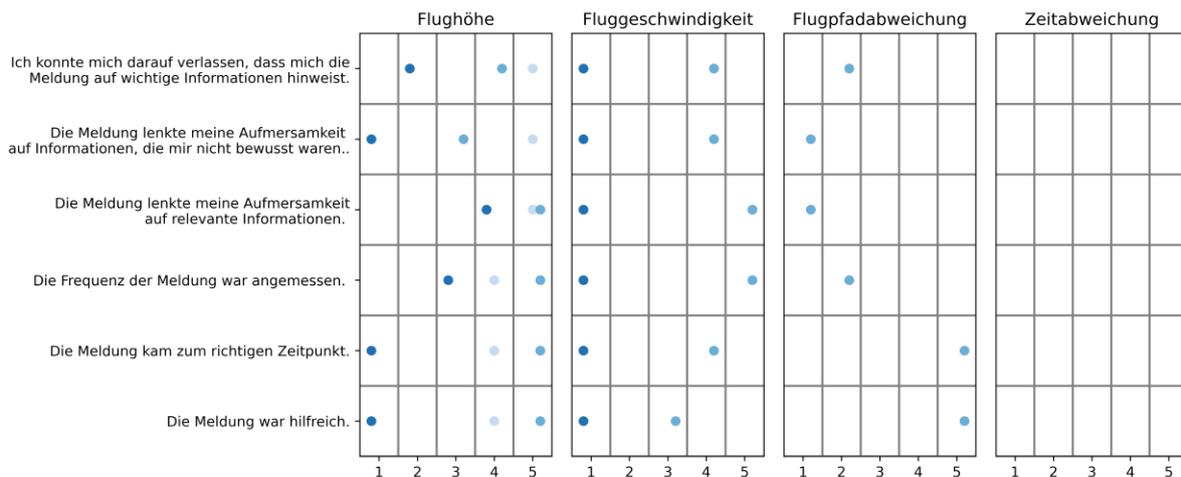


Abbildung 5-10 - Bewertung der Anwendungsfälle Flughöhe, Fluggeschwindigkeit, Flugpfadabweichung und Zeitabweichung

Pilot 3 gibt den Interventionen für beide Anwendungsfälle eine negative Bewertung. Bei der Flughöhemeldung wurde von Pilot 3 bemängelt, dass die Meldung im Sinkflug nicht notwendig war, da dieser vom Piloten eingeleitet wurde und er bereits am Erreichen der neuen Flughöhe arbeitete. In Bezug auf die Verbesserung der Flugeschwindigkeitsmeldung gab der Pilot folgenden schriftlichen Kommentar:

„DIE MELDUNG HAT MICH NUR DARAUF HINGEWIESEN, DASS ICH SCHON LANGE NICHT MEHR AUF DIE GESCHWINDIGKEITSANZEIGE GESCHAUT HABE. MEINE AUFMERSAMKEIT GALT JEDOCH DER ‚EARLY/LATE‘-FUNKTION, DIE EBENFALLS GESCHWINDIGKEITSABHÄNGIG IST. [...] IM LANDEANFLUG KONNTE DIE MELDUNG MICH AN ABWEICHUNGEN ZUR LANDEGESCHWINDIGKEIT ERINNERN.“

Dies zeigt, dass bei der Entwicklung des SB-adaptiven Assistenzsystems auf Redundanzen im Bereich der Anzeige geachtet werden muss. Wenn zwei Anzeigen verwandte Informationen auf unterschiedliche Arten enkodieren wie beispielsweise *Airspeed*- und *Early/Late-Indicator*, muss dies in der Interventionsgenerierung beachtet werden. Pilot 2 war der einzige, bei dem die Flugpfadabweichung aktiv war. Er selbst hat diese jedoch zu spät bemerkt, da er mit der Programmierung des Autopiloten beschäftigt war. Aus diesem Grund wurde zwar Zeitpunkt und Nützlichkeit gut bewertet, jedoch nicht die Salienz der Meldung.

Abbildung 5-11 zeigt die Bewertung der Anwendungsfälle „Nachricht“, „Autopilot“ und „FMS“. Die Meldung über eine Nachricht wurde von den zwei Piloten, die die Intervention bewerteten, in allen Kategorien als durchweg positiv bewertet. Besonders positiv hervorgehoben hat Pilot 2, dass der Inhalt der Nachricht vorgelesen wurde.

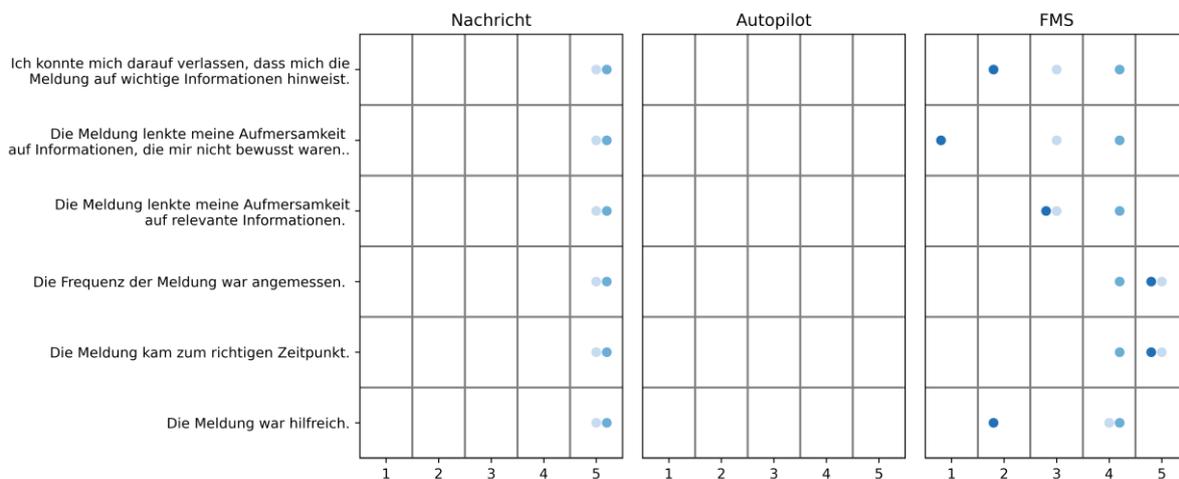


Abbildung 5-11 - Bewertung der Anwendungsfälle Nachricht, Autopilot und FMS

Die Autopilot-Nachricht wurde von keinem Piloten bewertet. Meldungen bezüglich des FMS wurde dagegen uneinheitlich bewertet. Zeitpunkt und Frequenz waren angemessen und richtig, jedoch wurde die Meldung nicht durchweg als relevant oder hilfreich bewertet. Pilot 3, der die Meldung als wenig hilfreich betrachtete, gab noch folgenden Verbesserungsvorschlag:

„WENN [VOM ASSISTENZSYSTEM] BEOBACHTET WIRD, DASS EIN KURVENFLUG EINGELEITET WURDE, DANN WIRD DIESE INFORMATION NICHT ZWANGSLÄUFIG BENÖTIGT.“

Dies zeigt sich eine Schwäche der SB-Messung, da diese nicht in der Lage ist, das Verständnis und die Projektionsfähigkeit des Piloten direkt zu messen. Ein erfahrener Pilot versteht, dass eine Änderung des Flugzeugkurses durch das FMS durch einen neuen Wegpunkt zustande kam.

Für die Anwendungsfälle neuer Bodenziele oder bodengestützter Bedrohung zeigt Abbildung 5-12 die Bewertung der Probanden. Hier zeigt sich ein positiveres Bild als bei den auditiven Nachrichten. Bis auf den Pilot 3, der die Information über ein neues Bodenziel nicht immer relevant fand, sind die Antworten in allen Dimensionen positiv.

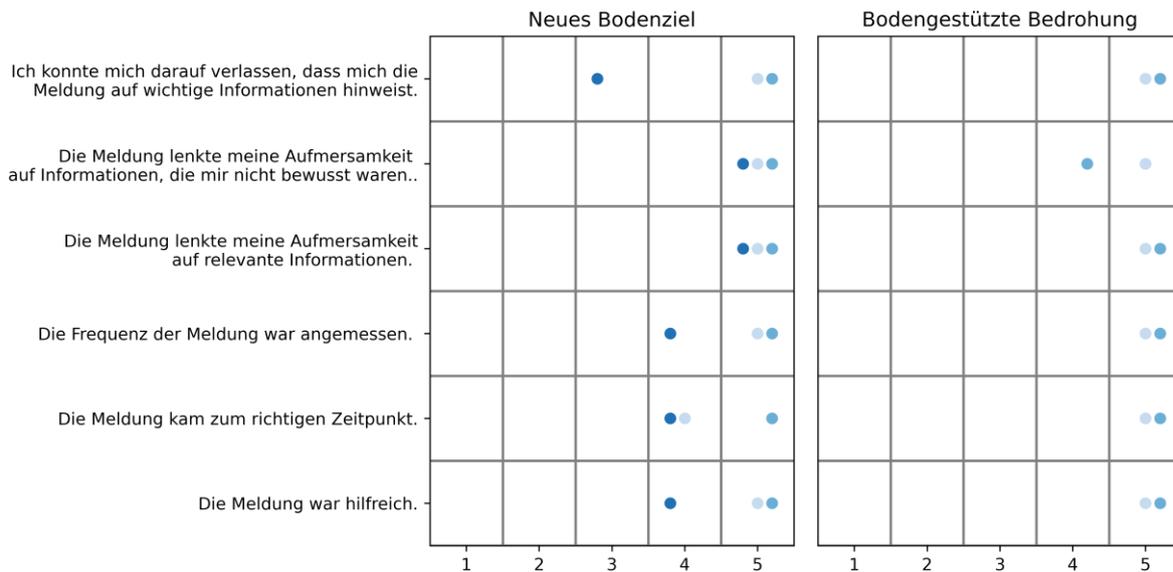


Abbildung 5-12 - Bewertung der Anwendungsfälle Neues Bodenziel und Bodengestützte Bedrohung.

Im Zusammenhang mit dem Bodenziel wurde zweimal der Wunsch einer auditiven Nachricht zusätzlich zur Visualisierung geäußert. Bei der Bedrohung wurde dieser Wunsch ebenfalls geäußert und dafür plädiert, eine auditive Unterscheidung der Benachrichtigung zwischen luft- gegenüber bodengestützten Bedrohung umzusetzen. Hier zeigt sich, dass die Bedrohungsinformationen stets als höchstrelevante Informationen gesehen werden, weshalb eine SB-adaptive Benachrichtigung als positiv gesehen wird. Offen ist aber, ob das adaptive Benachrichtigen gegenüber einem nicht-adaptiven Alarm die gleichen Ergebnisse erzielen könnte.

Abbildung 5-13 zeigt die weiteren Ergebnisse im Zusammenhang mit luftgestützter Bedrohung, Waffenanzahl und Waffeneinsatz. Die Benachrichtigung über luftgestützte Bedrohungen wurde – wie die bodengestützte Bedrohung – positiv bewertet. Zwei der Probanden befanden auch, dass diese Benachrichtigung auditiv oder auf anderen Bildschirmen sichtbar sein muss. Weniger eindeutig ist die Meinung über die niedrige Waffenzahl, da die Information den Piloten wohl bewusst war (siehe zweite Frage), jedoch von allen Probanden als nützlich eingeschätzt wurde. Der Benachrichtigung über den Einsatz von Waffen war für Pilot 1 nicht immer notwendig und damit nur moderat nützlich. Pilot 1 gab auch die Anregung, der Intervention mehr Missionsdaten hinzuzufügen wie beispielsweise Zeitpunkt der Wirkung oder des Ausbrennens eines Flugkörpers.

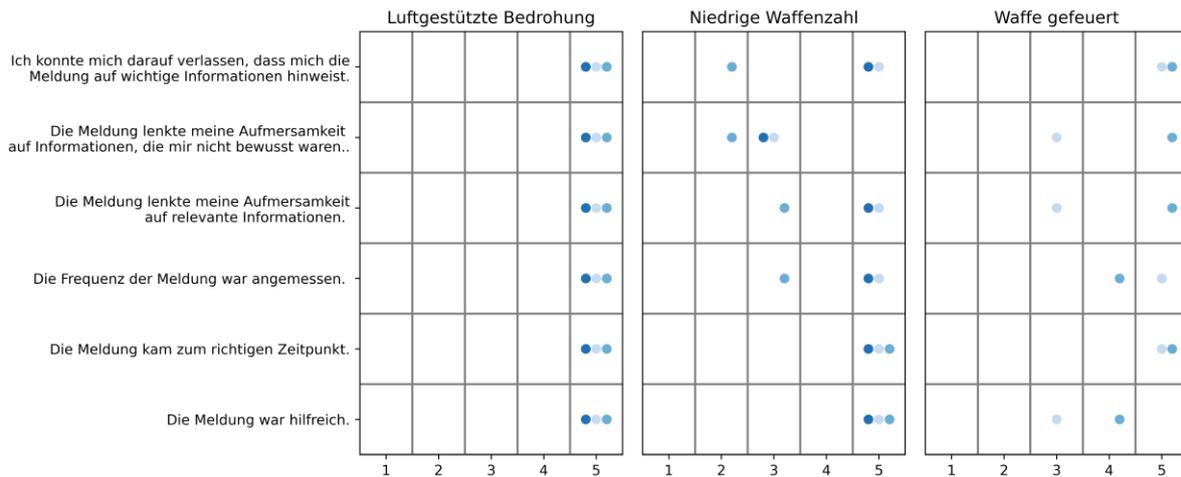


Abbildung 5-13 - Bewertung der Anwendungsfälle Luftgestützte Bedrohung, Niedrige Waffenzahl und Waffe gefeuert

### Gesamtsystembewertung

Ein Vergleich der Gesamtbewertungen zwischen Missionen und Piloten in Abbildung 5-14 zeigt, dass die Zuverlässigkeit des Gesamtsystems hoch eingeschätzt wurde.

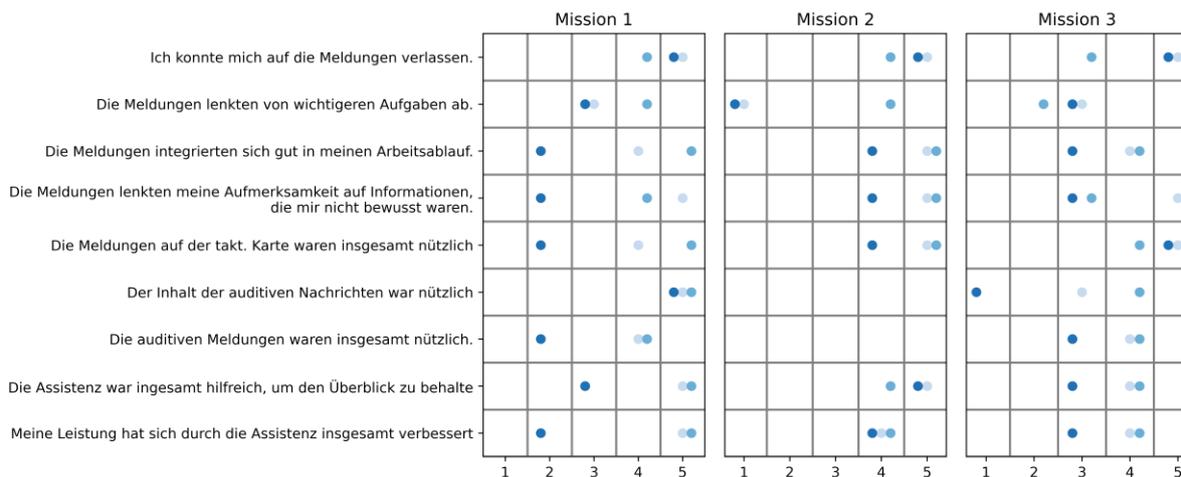


Abbildung 5-14 - Bewertung des Gesamtsystems über die drei Missionen

Der Vergleich der Missionen 1 & 2 im Zusammenhang mit der Ablenkung durch die Benachrichtigungen (Frage 2) zeigt, dass die auditiven Nachrichten als störend und von wichtigen Aufgaben ablenkend empfunden wurden. In den Antworten auf die Fragen 5-7 ist ersichtlich, dass Meldungen auf der taktischen Karte insgesamt positiver eingestuft werden als die auditiven Benachrichtigungen. Zwar wurde der Inhalt der auditiven Nachrichten als grundsätzlich sinnvoll eingeschätzt aber nicht als durchweg nützlich. Dies könnte an der disruptiven Natur der auditiven Meldungen liegen, da eine visuelle Benachrichtigung auch nur sichtbar ist, wenn die Aufmerksamkeit schon auf der taktischen Karte liegt, während die auditiven Nachrichten stets unterbrechen können. In Mission 2 und 3 wurde das System als leistungssteigernd bewertet, während Pilot 3 in der ersten Mission sich vom System gestört fühlte. Diese Erklärung wird auch durch den schriftlichen Kommentar von Pilot 3 in der dritten Mission bestätigt:

„DIE MELDUNGEN AUF DEM MHDD WAREN SINNVOLL, HINGEGEN HABEN DIE SPRACHMELDUNGEN EHER ABGELENKT UND WAREN NICHT BESONDERS HILFREICH.“

Ein weiterer Verbesserungsvorschlag von Pilot 2 war die bessere Priorisierung und Kategorisierung der Benachrichtigungen. In der Implementierung waren beispielsweise die Benachrichtigungen über Bedrohungen gleich gestaltet, wie die über eine neues Bodenziel, was nicht als adäquat empfunden wurde.

### 5.3.2.2 Situationsanalyse

Im Folgenden sollen einige Situationen aus den Versuchen herausgegriffen werden, die ein typisches Verhalten des adaptiven Systems darstellen. Dabei wird zwischen nützlichen, störenden beziehungsweise ignorierten und falsch-positiven Interventionen unterschieden.

#### Niedriges SB durch Unaufmerksamkeit

In einigen Situationen konnte das SB-adaptive System nützliche Interventionen generieren, welche anhand von zwei Beispielen im Folgenden demonstriert werden sollen.

Die erste Situation wurde in der ersten Mission mit dem ersten Probanden zum Simulationszeitpunkt 12:02:11 beobachtet. Vergleichbare Situationen konnten bei allen Probanden beobachtet werden. Der Proband merkte nicht, dass eine Nachricht mit einer neuen Flughöhe eintraf. Die SB-Abschätzung maß die Abweichung im assoziierten Informationsknoten und generierte einen Hinweis auf die neue Nachricht, welche direkt im Anschluss bemerkt und die neue Zielhöhe angesteuert wurde.

Die zweite Situation wurde beim ersten Probanden in der zweiten Mission zum Zeitpunkt 12:09:58 beobachtet und ist in den beiden Darstellungen Abbildung 5-15 und Abbildung 5-16 dargestellt.

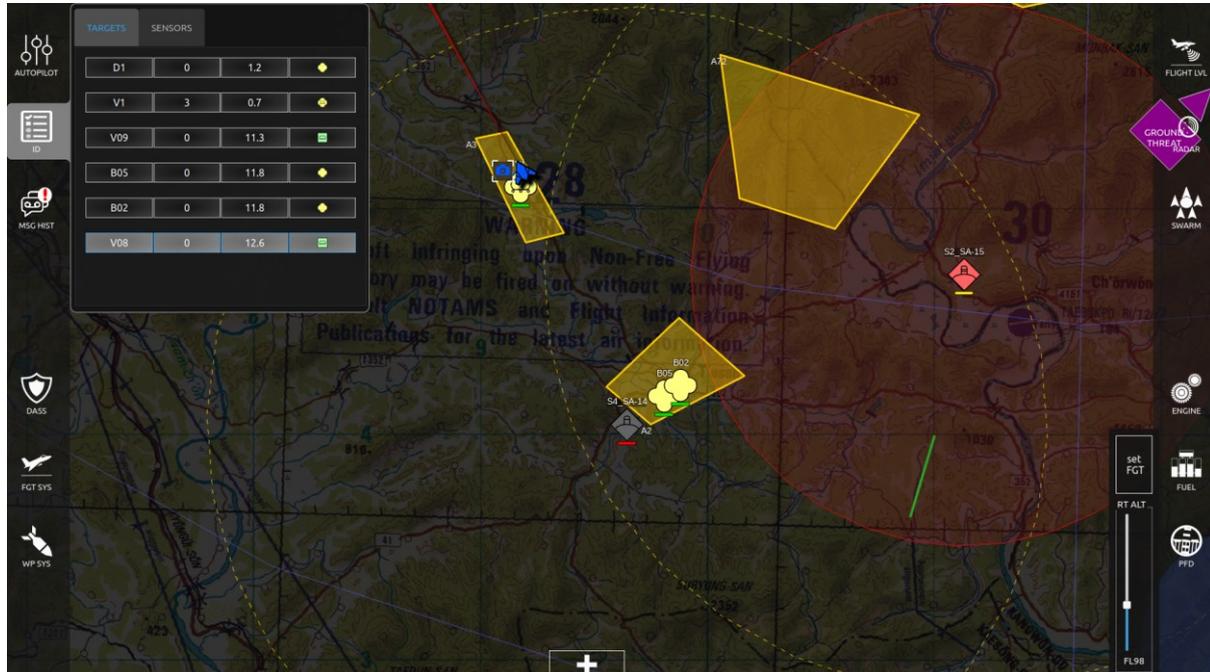


Abbildung 5-15 - Intervention (oben rechts) mit Hinweis auf bodengestützte Bedrohung, während der Proband auf einem kleinen Ausschnitt der Karte die Klassifikation einer SAM-Site durchführt

Der Proband war zunächst damit beschäftigt, die Klassifikation einer SAM-Site durchzuführen. In diesem Moment ist an einer Stelle außerhalb des Kartenausschnitts eine SAM-Site aufgetaucht, welche eine UCAV bedroht. Die Intervention durch das Assistenzsystem machte den Probanden darauf

aufmerksam, welcher direkt darauf reagierte. Ähnliche Situationen konnten vor allem im Zusammenhang mit unbekanntem oder dynamischen Objekten auf der taktischen Karte beobachtet werden.

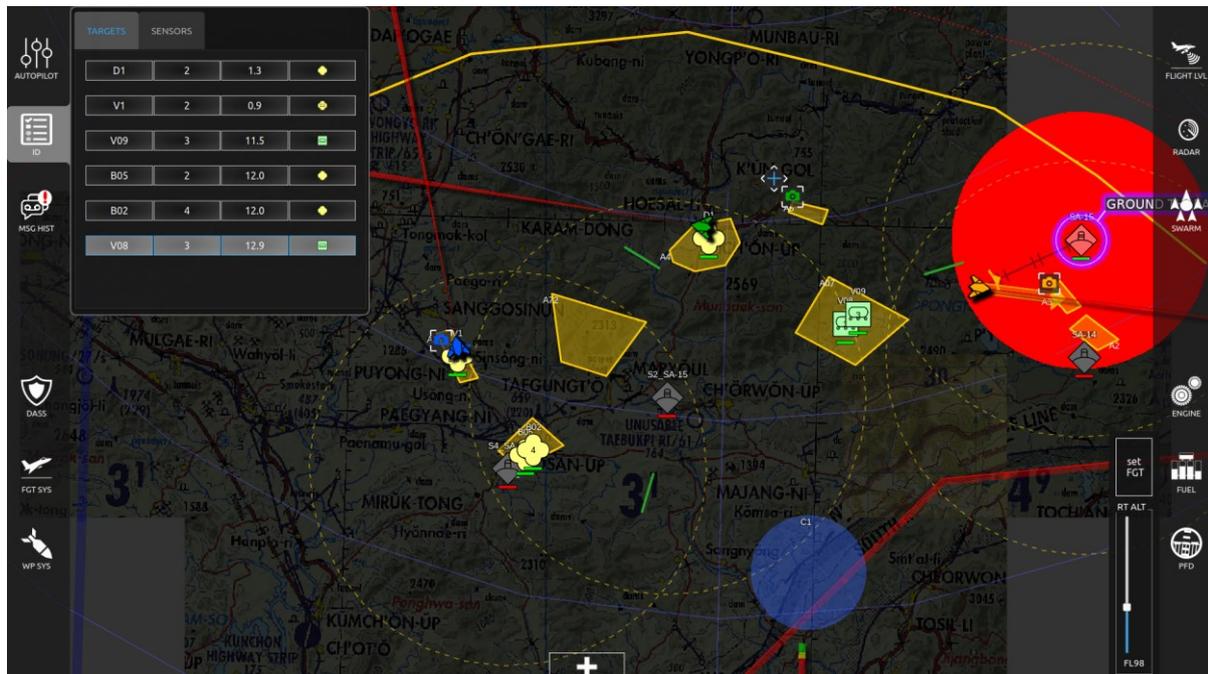


Abbildung 5-16 - Reaktion des Probanden auf Intervention durch Anpassung des Kartenausschnitts und Planung der Bekämpfungsaufgabe

### Störung durch Meldungen in Situationen hoher Arbeitsbelastung

Eine problematische Eigenschaft des SB-adaptiven Systems trat in Situationen hoher Belastung auf, was anhand von zwei Situationen beschrieben werden soll. Die erste dieser beiden Situationen trat beim zweiten Probanden in der ersten Mission auf. Dort hatte der Proband im Zeitraum 12:01:15 – 12:02:00 Probleme bei der Bedienung des Autopiloten. Die Probleme rührten von der Eigenschaft des Autopiloten sich bei Bewegung des Steuerknüppels zu deaktivieren, was den Probanden wiederholt verwirrte. Während dieses Zeitraums ignorierte der Proband seine eigentliche Aufgabe, die Route zu halten und verließ den geplanten Flugpfad. Die SB-Abschätzung erkannte, dass dem Piloten diese Abweichung vom Flugpfad nicht bewusst war und generierte folglich nach einer Zeit von 10 Sekunden eine Intervention auf der taktischen Karte und eine auditive Benachrichtigung. Diese wurde vom Piloten ignoriert, was der Proband im Nachhinein auf seine hohe Belastung durch die Bedienung des Autopiloten erklärte.

Die zweite dieser Situationen trat beim zweiten Probanden in der zweiten Mission auf und ist in Abbildung 5-17 dargestellt. Es lag eine Situation hoher Aufgabenlast vor, da zwei SAM-Sites gerade aktiv geworden sind und gleichzeitig ein feindliches Flugzeug in den Luftraum eingedrungen ist. Der Proband konzentriert sich in dieser Situation auf die dringendste Aufgabe; die Reaktion auf den Beschuss des grünen UCAVs. Um dieser priorisierten Aufgabe gerecht zu werden, vernachlässigt der Proband andere Informationsquellen. Das führt zu mehr Interventionen durch das Assistenzsystem und damit zu einer zusätzlichen Belastung des Probanden. Hier wäre eine Priorisierung der Informationen notwendig, um die störenden Benachrichtigungen zu filtern beziehungsweise zu verzögern.

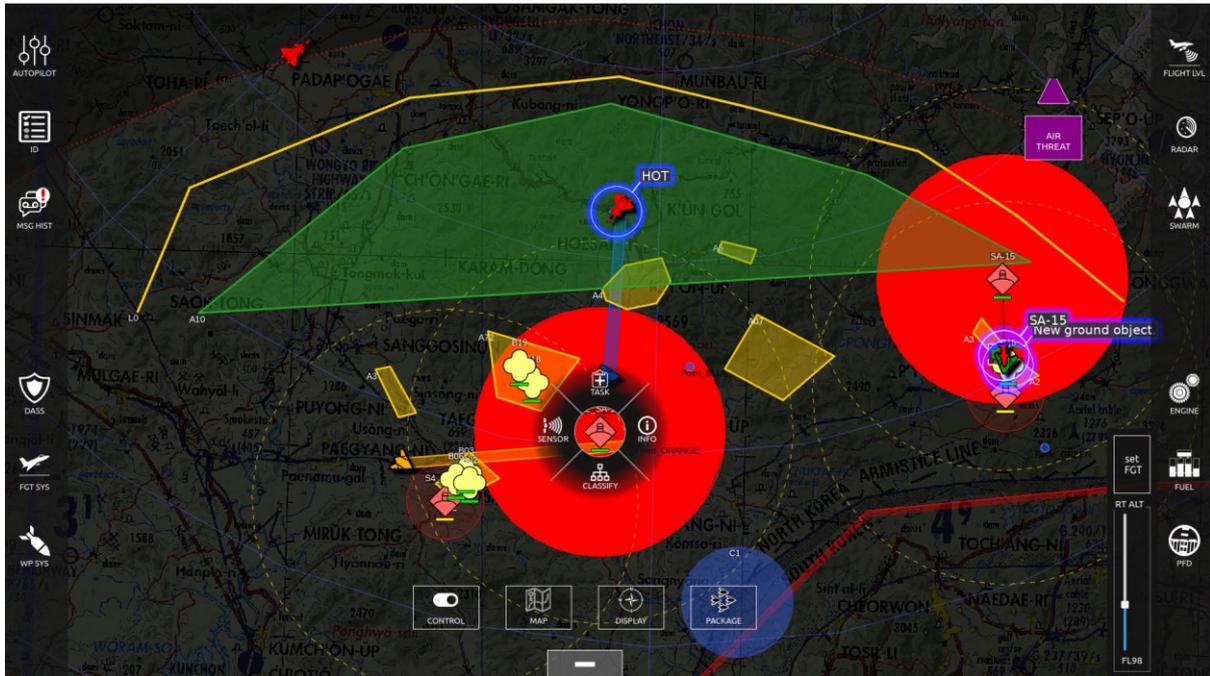


Abbildung 5-17 - Darstellung des MHDD zum Zeitpunkt 12:08:15 bei Proband 2 mit vier aktiven Benachrichtigungen

### Falsch-positive Benachrichtigung

Neben den nützlichen Situationen gab es auch Benachrichtigungen, die zwar inhaltlich korrekt aber dem Piloten bereits bewusst waren. Beispielsweise trat in der ersten Mission die Flughöhenbenachrichtigung bei zwei Piloten während einem Steig- beziehungsweise Sinkflug auf, was die Piloten im Nachgang als überflüssig bewerteten, da die Benachrichtigung einerseits nur Flughöhen meldete, die aufgrund der starken Höhenänderung zum Zeitpunkt der Meldung bereits wieder veraltet waren, und andererseits den Piloten ja bewusst ist, dass sich ihre Höhe gerade ändert, auch wenn sie diese Änderung nicht mit hoher Frequenz überwachen. Ein konkretes Beispiel gab es bei Pilot 1 zur Missionszeit 12:03:27 (T=147s), welcher in der Abbildung 5-18 dargestellt ist.

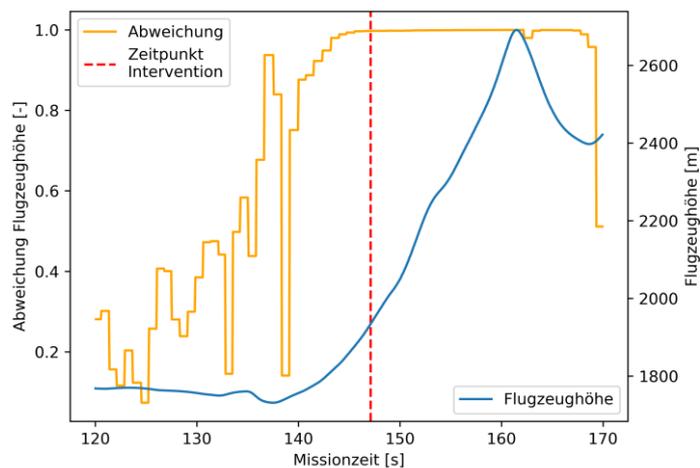


Abbildung 5-18 - Verlauf von Flughöhe (blau) und Abweichung im Knoten "Flugzeughöhe" (orange). In rot ist der Zeitpunkt der Intervention dargestellt.

Zu diesem Zeitpunkt wurde ein Steigflug zu einer neuen gewünschten Flughöhe eingeleitet, weshalb sich die Flughöhe änderte. Der Pilot fixierte die Flughöheninformation zu Beginn (T=139s) des Steigflugs und fixierte im Anschluss nur die Außensicht, ohne wieder auf den Höhenmesser zu blicken. Von der SB-Abschätzung wird dies fälschlicherweise als fehlendes Bewusstsein für die Höhe interpretiert und eine Intervention ausgelöst. Dieses Beispiel zeigt eine weitere Schwäche der SB-Bewertung, welche bisher nicht in der Lage ist, mentale Modelle oder Projektionen des Piloten zu messen.

### 5.3.3 DISKUSSION

Zunächst sollte der Studie geklärt werden, ob die Auswahl der Anwendungsfälle sinnvoll und die Informationen im Aufgabenkontext eines Kampfflugzeugpiloten relevant sind. Die Ergebnisse sprechen stark dafür, dass dies der Fall ist und die Benachrichtigungen inhaltlich relevant sind. Dies lässt sich vor allem an den Antworten auf die Frage „Die Meldung lenkte meine Aufmerksamkeit auf relevante Informationen“ ablesen, welche fast ausschließlich positiv waren. Die Piloten steuerten zudem wenig andere Anwendungsfälle bei, die ihnen bei den Missionen gefehlt haben, was aber auch an der Ausrichtung der Aufgabentypen auf die Interventionen liegen kann.

Die zweite Forschungsfrage, ob die SB-adaptive Intervention in Kontext komplexer Missionen nützlich ist, kann nicht ganz so eindeutig bestätigt werden. In einigen Fällen wurde der Pilot auf Informationen hingewiesen, die ihm nicht bewusst waren, weil er die Überwachung vernachlässigt hatte. Die Bewertung durch die Piloten zeigt grundsätzlich, dass die Benachrichtigungen auf der taktischen Karte nützlicher sind als die auditiven Benachrichtigungen. Zwei Eigenschaften der Interventionsgestaltung könnten hierfür ausschlaggebend sein: Der bereits diskutierte disruptive Charakter der auditiven Nachrichten in Kombination mit der Auswahl des richtigen Zeitpunkts einer Intervention. Die Interventionsgenerierung sieht hierfür eine Zeitkonstante vor, die abgewartet wird, bevor bei hoher Abweichung eine Intervention ausgelöst wird. Dieses Vorgehen ist nachteilhaft im Zusammenhang mit belastenden Arbeitssituationen. Diese Situationen zeichnen sich dadurch aus, dass Aufgaben priorisiert werden müssen und dass Informationsquellen vernachlässigt werden müssen, um sich auf eine Aufgabe zu konzentrieren. Genau diese Vernachlässigung kann die SB-Abschätzung messen, zieht bei der Interventionsgenerierung aber die falschen Schlüsse, indem es saliente Benachrichtigungen erstellt und den Piloten bei seiner Arbeit stört. Hier müsste das System erkennen können, dass der Pilot momentan ungestört arbeiten sollte. Ein zweiter Grund für die bessere Bewertung der visuellen Benachrichtigungen ist auch, dass der Inhalt sich fast ausschließlich auf taktisch relevante Informationen (beispielsweise Bedrohungen) bezogen haben, während die auditiven Informationen meist Flugparameter oder Systemzustände zum Inhalt hatten, welche im Missionskontext weniger relevant sind.

Bei der Gestaltung der Studie wurde darauf geachtet, dass die Probanden den Mechanismus des Assistenzsystems nachvollziehen können. Das hatte gegenüber dem Versuch im Abschnitt 4.4 den Vorteil, dass die Piloten nicht durch die Adaptivität des Systems verwirrt wurden. Ein Pilot kommentierte im Nachgang des Versuchs sogar, dass er sich auf die Benachrichtigung über eine neue Nachricht verlassen hat und gar nicht auf die Anzeige blickte. Dies könnte als eine Ausprägung von Kooperation zwischen Piloten und Assistenzsystem interpretiert werden. Aus dieser Beobachtung heraus erscheint es auch plausibel, dass ein Pilot den Umgang mit einem adaptiven System erlernen und einen Modus der Zusammenarbeit finden muss, um das volle Potenzial des Mensch-Maschine-Systems auszuschöpfen.

## Limitierungen

Der Versuch hat zwei Limitierungen. Durch die wenigen Versuchsteilnehmer haben die Ergebnisse keine quantitative Aussagekraft, sondern basieren auf der qualitativen Nutzerbewertung sowie Beobachtungen des adaptiven Verhaltens im Versuch. Ferner kann nicht gesichert gesagt werden, dass die Adaptivität auf Basis der SB-Abschätzung einen wesentlichen Vorteil gegenüber einem Benachrichtigungssystem hat, das nicht auf Abschätzung des kognitiven Zustands basiert.

## Zusammenfassung

Die zentralen Erkenntnisse dieses Versuchs sind die Folgenden:

- Die GDTA lieferte plausible und relevante Anwendungsfälle und eignet sich grundsätzlich als Methode zur Wissensgewinnung bei der Gestaltung der SB-adaptiven Assistenz
- Der größte Nutzen der SB-Assistenz liegt in der Benachrichtigung über nicht bemerkte, taktisch relevante Informationen mit hoher Priorität
- Bei der Gestaltung der Interventionen muss darauf geachtet werden, dass die Benachrichtigungen in Situationen hoher Aufgabenlast nicht stören.
- Wenn Bediener die Funktionsweise der Adaptivität verstehen, können sie sich das zu Nutze machen.

## 6 ZUSAMMENFASSUNG & AUSBLICK

In der vorliegenden Arbeit wurde ein Messsystem zur kontinuierlichen Abschätzung von Aufmerksamkeit und SB eines Bedieners vorgestellt. Das Messsystem wurde in einen Cockpitsimulator integriert und in verschiedenen Aufgabentypen validiert. Darauf aufbauend wurde ein adaptive Interventionsgenerierung entwickelt, welche die SB-Abschätzung nutzt, um einen Bediener auf ihm nicht bewusste Informationen aufmerksam zu machen. Diese Interventionsgenerierung wurde zunächst in Validierungsstudien erprobt und dann in die Anwendung von MUM-T für Kampfflugzeuge transferiert. Dieser Transfer wurde abschließend von Experten mittels virtueller Cockpitsimulation von militärischen Missionen bewertet.

### 6.1 BEITRÄGE DER ARBEIT

Die Beiträge dieser Arbeit liegen in verschiedenen Bereichen, die in der Theorie von SB verortet werden können, was in der Abbildung 6-1 visualisiert wird. Die Darstellung baut auf dem Modellschaubild von Situationsbewusstsein (siehe Abbildung 2-6 ) in Abschnitt 2.2.1.1 auf. Sie zeigt das Assistenzsystem als Agenten und ordnet die Beiträge dieser Arbeit den verschiedenen Komponenten von Situationsbewusstsein zu.

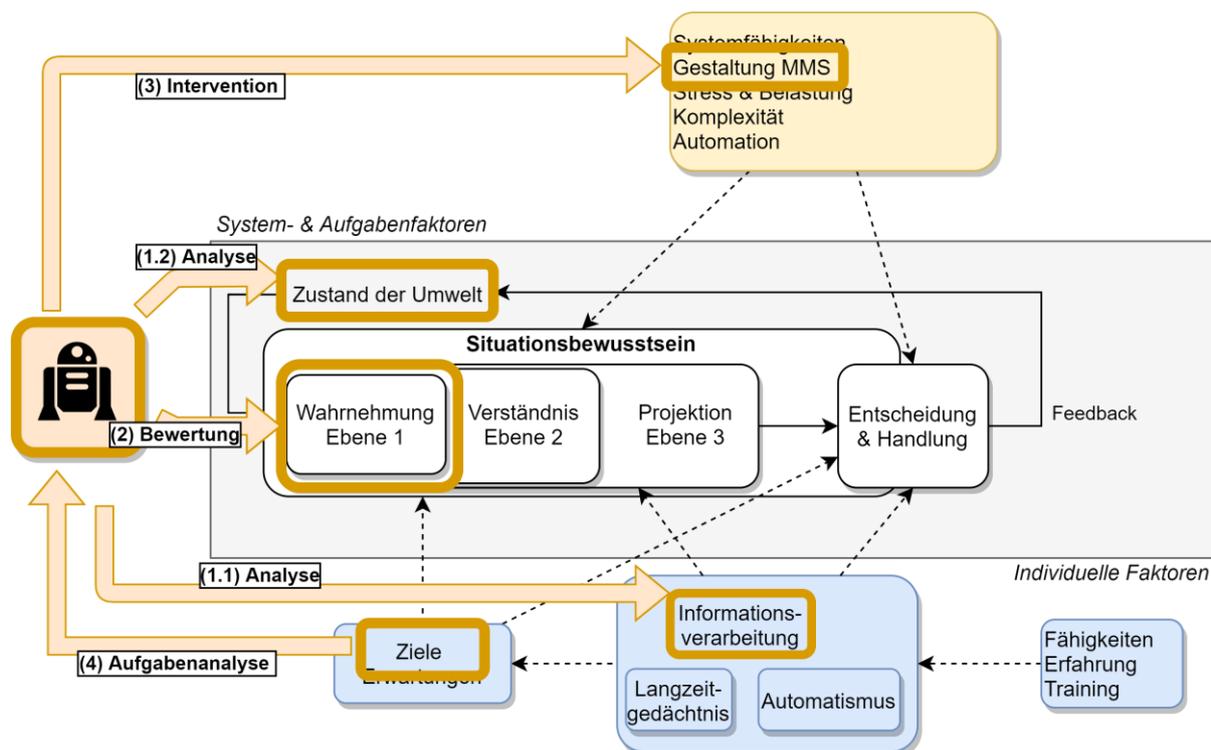


Abbildung 6-1 - Assistenzsystem (Agent, links) mit Zusammenhängen zum Modell von SB nach Endsley und Jones (2012)

Ein erster Beitrag ist das vorgestellte **Verfahren zur SB-Bewertung basierend auf Blickbewegungs- und Interaktionsanalyse** (siehe (1.1), (1.2) und (2) in Abbildung 6-1). Hierbei wurde auf bestehende Verfahren aufgebaut und durch das generische SB-Modell mit informationspezifischer Modellierung, Relationen und datentyp-abhängiger Abweichungsberechnung eine Möglichkeit geschaffen, die hier validierten Ansätze auf andere Domänen zu übertragen. Dies trägt zu einer besseren Übertragbarkeit der situativen Messmethode bei (vergleiche Tabelle 2-1). Es konnte darüber hinaus nachgewiesen

werden, dass das konzeptionierte SB-Messsystem in der Lage ist, fehlendes oder falsches SB informationsspezifisch zu messen. Hier wurden sowohl Korrelationen mittlerer Maße mit der Bedienerleistung gezeigt als auch individuelle Situationen, in denen der Grund für Fehler durch die SB-Bewertung abgebildet wurde. Jedoch zeigte sich auch, dass der Ansatz anfällig für falsch-positive Messungen ist. Vor allem bei der Messung in Anzeigeoberflächen mit hoher Informationsdichte (bspw. taktische Karte) konnte das System nicht robust feststellen, welche der vielen Informationen beachtet wurde. Außerdem hat sich in den Studien gezeigt, dass die SB-Bewertung vorwiegend für die erste Ebene (Wahrnehmung) funktioniert, was in der Abbildung grafisch angedeutet ist.

Der zweite Beitrag ist **die Evaluierung eines realisierten SB-adaptives System**, das durch Anpassung an die SB-Bewertung dem Bediener Informationen mitteilt, die ihm nicht bewusst sind (siehe 3 in Abbildung 6-1). Die Gestaltung des MMS wird im klassischen Modell von Endsley und Jones (2012) noch als statischer Systemfaktor verstanden und wird mithilfe des adaptiven Ansatzes dynamisch. Die Interventionsstrategie konnte die Leistung für sensor-motorische Flugführungs- sowie Erkennungsaufgaben steigern. Vor allem im Kontext von unerwarteten Systemwertänderungen zeigte sich die große Stärke des Systems. Eine Schwäche des Systems war die fehlende Modellierung für priorisierte Aufgaben und die Tendenz, bei vorübergehender Fokussierung des Bedieners auf eine Aufgabe, die Arbeitsbelastung durch zusätzliche Interventionen zu erhöhen. Die subjektive Bewertung des Systems fiel dadurch gemischt aus. Zwar waren die Anwendungsfälle sinnvoll und wurden im Falle von kritischen Informationen stets positiv bewertet. Jedoch führten falsch-positive Meldungen zu Verlust von Vertrauen in das System oder wurden als störend empfunden. Das Verständnis über die Funktionsweise des Systems spielt hier eine große Rolle.

Der dritte Beitrag dieser Arbeit ist eine Untersuchung, **welche Arten von Fehlern mit einem SB-adaptiven Ansatz vermieden werden können**. Da die SB-Bewertung auf der ersten Ebene des Modells am genauesten ist, können auch die Fehlerarten dem in Abschnitt 2.3.1 beschriebenen Fehlermodell mit Einschränkungen entnommen werden. Die Tabelle 6-1 führt die Fehlerarten auf und beschreibt, ob eine Vermeidung durch die SB-adaptive Assistenz möglich ist.

Fehlerart (Ebene 1)	Vermeidung durch SB-adaptive Assistenz
Information nicht verfügbar	Nicht möglich, da keine Abweichung berechnet werden kann, wenn das System die Information nicht kennt.
Information nicht bemerkt	Möglich und nachgewiesen
Information nicht überwacht	Möglich und nachgewiesen
Information falsch wahrgenommen	Nicht möglich, da Fehler der Wahrnehmung nicht gemessen werden kann
Information vergessen	Nicht möglich, da SB-Modell das „Vergessen“ nicht modelliert.

Tabelle 6-1 - Überblick über Fehlerarten der ersten Ebene und eine mögliche Vermeidung durch SB-adaptive Assistenz.

Der vierte Beitrag ist **eine Methode zur Realisierung des SB-adaptiven Assistenzsystem in einer komplexen Aufgabenumgebung** (siehe 4 in Abbildung 6-1). Hierfür wurde eine Methode vorgestellt, um das SB-adaptive Assistenzsystem auf Basis einer Aufgabenanalyse zu entwickeln. Die Aufgabenanalyse bietet eine Möglichkeit, die Ziele von Operateuren und relevante Informationen zu identifizieren. Dies ist für Entwickler von adaptiven Assistenzsystemen in komplexen Aufgabenumgebungen eine wertvolle Hilfestellung.

## 6.2 VERBESSERUNGSPOTENZIALE

Die Abbildung 6-1 zeigt, dass einige Bereich der Situationsbewusstseinstheorie nicht abgebildet wurden während andere abgebildet wurden, aber noch Potenzial für Verbesserungen bieten.

### Individuelle Faktoren

Einer der nicht betrachteten individuellen Faktoren ist die Erfahrung und das Training, das die Bediener mitbringen. Der Erfahrungsstand wurde nicht explizit modelliert und es wurde vereinfachend angenommen, dass die Bedienergruppe professioneller Kampfflugzeugpiloten homogen in Kenntnis und Erfahrung ist. Eine interessante Frage ist hier, ob es Unterschiede zwischen Novizen und Profis bei der SB-Abweichung gibt und wie diese aussehen.

### Analyse von Aufmerksamkeit

Bei der Analyse der Aufmerksamkeit gibt es Möglichkeiten zur Erhöhung der Robustheit der Wahrnehmungserkennung durch Integration von Modellen der räumlichen Salienz, also einer Analyse der „Wahrnehmbarkeit“ eines Objekts und der damit verbundenen Wahrscheinlichkeit, dass es bei einer Fixation auch wahrgenommen wurde. Eine Kombination mit anderen Messmethoden könnte genutzt werden, um unaufmerksame Zustände, wie bspw. Müdigkeit, zu erkennen und integrieren.

### Situationsbewusstseinsbewertung

Bei der Gestaltung des SB-Modells ist eine mögliche Weiterentwicklung, dass Bedienerziele und Aufgabeninformationen direkt im Modell abgebildet werden und nicht erst im Nachgang durch die Interventionsgenerierung beachtet werden. Dies ließe die Möglichkeit zu, aus dem Modell direkt alle relevanten Informationen auszulesen und würde den Interventionsmechanismus unabhängig von den Kontextinformationen der Aufgabensituation machen. Voraussetzung ist dann wiederum, dass das SB-Modell mit diesen Kontextinformationen versorgt wird.

Wie aus der Tabelle 6-1 hervorgeht, könnte durch Modellierung von „Vergessen“ von Informationen eine weitere Fehlerkategorie durch das System abgedeckt werden. Im Zusammenhang mit dynamischen Systemen, sollte man sich zunächst auf die Modellierung des Arbeitsgedächtnisses konzentrieren. Hierfür könnte auf Vorarbeiten aus der Psychologie wie die Studie von Oberauer (2019) zurückgegriffen werden.

Schließlich ist Abschätzung von höheren Ebenen von Situationsbewusstsein eine bleibende Herausforderung. Hierfür könnte eine Modellierung der mentalen Modelle der Bediener helfen, womit dann vorhergesagt werden kann, ob ein Bediener sich einer Änderung bewusst ist. Die Frage, welches mentale Modell ein Bediener verwendet, ist jedoch nicht trivial.

### Interventionsgestaltung

Bei der Interventionsgestaltung gibt es Verbesserungsmöglichkeiten bei der Wahl des Interventionszeitpunkts und der -modalität. Wie die Ergebnisse des letzten Versuchs zeigen, hat dies einen großen Einfluss darauf, wie die Intervention auf den Bediener wirkt. Ein Ansatz wäre hier, eine Tätigkeitserkennung zu integrieren, welche den Arbeitsfokus und -fluss erkennt, wann ein guter Zeitpunkt zur Intervention wäre und wie diese Intervention am besten aussähe. Ersteres ist Forschungsgegenstand von aktuellen Studien am IFS der UniBwM. Letzteres könnte beispielsweise nach dem Konzept von Maiwald und Schulte (2012) umgesetzt werden. Dies entspricht dann einer Integration des Systemfaktors „Stress & Arbeitsbelastung“ (siehe Abbildung 6-1).

Eine Weiterentwicklung der Interventionsarten ist, nicht nur Benachrichtigungen zu generieren, sondern auch eine Anpassung der Automationsgrade basierend auf der SB-Abschätzung vorzunehmen, vergleichbar mit einigen Studien der beanspruchungsadaptiven Interventionsgenerierung (Brand & Schulte, 2021; Kaber & Endsley, 2004). Dies würde einer Berücksichtigung des Systemfaktors „Automation“ entsprechen (siehe Abbildung 6-1). Hier sind verschiedene Strategien denkbar, wie beispielsweise die Erhöhung des Automationsgrad in Aufgaben, in denen ein Bediener ein niedriges oder hohes SB besitzt. Beide Strategien sind plausibel, denn die erste Strategie entlastet den Bediener, während die zweite den Bediener in Aufgaben einbezieht, in denen er sich nicht auskennt, um *Out-of-the-loop*-Phänomene zu verhindern.

### 6.3 AUSBLICK

Das vorgestellte System könnte nach weiterer Verbesserung in vielen Anwendungen zum Einsatz kommen und ist nicht auf das Flugzeugcockpit beschränkt. Alle Arbeitssituationen, die durch eine Vielzahl dynamischer Informationen bestimmt sind, könnten davon profitieren. Abgesehen vom Einsatz in einem adaptiven Assistenzsystem, ist auch der Einsatz der SB-Messung in Trainingssystemen denkbar. Hier könnte das objektive Maß der SB-Abweichung beispielsweise als Qualitätskriterium für gutes Überwachungsverhalten eines Piloten dienen.

Die vorliegende Arbeit demonstriert, dass die maschinelle Beobachtung eines Bedieners und die Abschätzung von kognitiven Zuständen geeignet ist, um Kooperation zwischen Menschen und Maschine zu ermöglichen. Voraussetzung für eine Leistungssteigerung des Mensch-Maschine-Systems ist eine robuste Bewertung des kognitiven Zustands sowie eine adäquate Gestaltung der Intervention. Adaptive Systeme stoßen bei professionellen Bedienern wie Piloten meist auf Skepsis, da ihr Verhalten als nicht transparent und vorhersehbar eingeschätzt wird. Um den Aufwand der Implementierung zu rechtfertigen und Nutzer zu überzeugen, muss ein adaptives System einen spür- und messbaren Vorteil durch Kooperation bieten. Hierfür muss ein Bediener das notwendige Vertrauen in das System gewinnen, damit sich ein kooperatives Verhältnis entfalten kann. Dieser Aspekt in der Interaktion von Menschen und Maschine erinnert an zwei kooperierende Menschen, deren Zusammenarbeit auch durch den Aufbau von Vertrauen gefördert wird und unter Missverständnissen leiden kann.

## A. ABKÜRZUNGSVERZEICHNIS

---

AA	Adaptive Assistenz
A/I	Air Interdiction
AoI	Area of Interest
ATPL	Airline Transport Pilot License
C2	Command-and-Control
EEG	Elektroenzephalographie
EKG	Elektrokardiografie
FCM	Fuzzy Cognitive Map
FMS	Flight Management System
fNIRS	Funktionelle Nahinfrarotspektroskopie
GDTA	Goal Directed Task Analysis
HUD	Head-Up Display
HOTAS	Hands-on-Throttle-and-Stick
IFS	Institut für Flugsysteme
LoA	Level of Automation
MHDD	Multi-functional Head-Down Display
MMS	Mensch-Maschine-Schnittstelle
MUM-T	Manned-Unmanned-Teaming
PFD	Primary Flight Display
SAGAT	Situation Awareness Global Assessment Technique
SART	Situation Awareness Rating Technique
S/A	Simulation/Automation
SB	Situationsbewusstsein
SEAD	Suppression of Enemy Air Defense
SHDD	Side Head-Down Display
SPAM	Situation Present Assessment Method
SPL	Sport Pilot Licence
SVS	Synthetic Visual System
TMG	Touring Motor Glider
TO	Tactical Object
UAV	Unmanned Aerial Vehicle
UCAV	Unmanned Combat Aerial Vehicle
UniBwM	Universität der Bundeswehr München

## B. LITERATURVERZEICHNIS

- Adams, M. J., Tenney, Y. J. & Pew, R. W. (1995). Situation Awareness and the Cognitive Management of Complex Systems. *Human Factors*, 37(1), 85–104.
- Airbus (Hrsg.). (2021). *FCAS: Future Combat Air System*. <https://www.airbus.com/en/products-services/defence/multi-domain-superiority/future-combat-air-system-fcas>
- Alexander, A. L., Wickens, C. D [Christopher D.] & Hardy, T. J. (2005). Synthetic vision systems: the effects of guidance symbology, display size, and field of view. *Human Factors*, 47(4), 693–707. <https://doi.org/10.1518/001872005775571005>
- Arapakis, I. & Leiva, L. A. (2020). Learning Efficient Representations of Mouse Movements to Predict User Attention. In J. Huang, Y. Chang, X. Cheng, J. Kamps, V. Murdock, J.-R. Wen & Y. Liu (Hrsg.), *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 1309–1318). ACM. <https://doi.org/10.1145/3397271.3401031>
- Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Bonelli, S., Golfetti, A., Pozzi, S., Imbert, J.-P., Granger, G., Benhacene, R. & Babiloni, F. (2016). Adaptive Automation Triggered by EEG-Based Mental Workload Index: A Passive Brain-Computer Interface Application in Realistic Air Traffic Control Environment. *Frontiers in human neuroscience*, 10, 539. <https://doi.org/10.3389/fnhum.2016.00539>
- Bainbridge, L. (1983). IRONIES OF AUTOMATION. In *Analysis, Design and Evaluation of Man-Machine Systems* (S. 129–135). Elsevier. <https://doi.org/10.1016/B978-0-08-029348-6.50026-9>
- Bakdash, J. Z., Marusich, L. R., Cox, K. R., Geuss, M. N., Zaroukian, E. G. & Morris, K. M. (2022). The validity of situation awareness for performance: a meta-analysis. *Theoretical Issues in Ergonomics Science*, 23(2), 221–244. <https://doi.org/10.1080/1463922X.2021.1921310>
- Belo, J., Clerc, M. & Schön, D. (2021). EEG-Based Auditory Attention Detection and Its Possible Future Applications for Passive BCI. *Frontiers in Computer Science*, 3, Artikel 661178. <https://doi.org/10.3389/fcomp.2021.661178>
- Bogler, C., Mehnert, J., Steinbrink, J. & Haynes, J.-D. (2014). Decoding vigilance with NIRS. *PloS one*, 9(7), e101729. <https://doi.org/10.1371/journal.pone.0101729>
- Bolstad, C. A., Riley, J. M., Jones, D. G. & Endsley, M. R. (2002). Using Goal Directed Task Analysis with Army Brigade Officer Teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46(3), 472–476. <https://doi.org/10.1177/154193120204600354>
- Bosse, T., van Lambalgen, R., van Maanen, P.-P. & Treur, J. (2009). Attention Manipulation for Naval Tactical Picture Compilation. *EEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, 2, 450–457. <https://doi.org/10.1109/WI-IAT.2009.194>
- Brams, S., Hooge, I. T. C., Ziv, G., Dauwe, S., Evens, K., Wolf, T. de, Levin, O., Wagemans, J. & Helsen, W. F. (2018). Does effective gaze behavior lead to enhanced performance in a complex error-detection cockpit task? *PloS one*, 13(11), e0207439. <https://doi.org/10.1371/journal.pone.0207439>
- Brand, Y. & Schulte, A. (2018). Design and Evaluation of a Workload-Adaptive Associate System for Cockpit Crews. In D. Harris (Hrsg.), *Engineering Psychology and Cognitive Ergonomics* (S. 3–18). Springer International Publishing.
- Brand, Y. & Schulte, A. (2021). Workload-adaptive and task-specific support for cockpit crews: design and evaluation of an adaptive associate system. *Human-Intelligent Systems Integration*. Vorab-Onlinepublikation. <https://doi.org/10.1007/s42454-020-00018-8>
- Broadbent, D. E. (1958). *Perception and Communication*.
- Bundeswehr Intranet (Hrsg.). (2020, 28. Juli). *Definition MUM-T (FW-0000043249)*.

- Calhoun, G. L. [G. L.], Ward, V. B. R. & Ruff, H. A. [H. A.] (2011). Performance-based Adaptive Automation for Supervisory Control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 55(1), 2059–2063. <https://doi.org/10.1177/1071181311551429>
- CMU Flite: Speech Synthesizer. (2018, 4. März). <http://cmuflite.org/>
- Colvin, K., Dodhia, R. & Dismukes, R. (2005). Is pilots' visual scanning adequate to avoid mid-air collisions? In *2005 International Symposium on Aviation Psychology*.
- Dehais, F., Peysakhovich, V., Scannella, S. [Sébastien], Fongue, J. & Gateau, T. (2015). "Automation Surprise" in Aviation. In B. Begole, J. Kim, K. Inkpen & W. Woo (Hrsg.), *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (S. 2525–2534). ACM. <https://doi.org/10.1145/2702123.2702521>
- Dehais, F., Roy, R. N. [R. N.] & Scannella, S. [S.] (2019). Inattentive deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural brain research*, 360, 51–59. <https://doi.org/10.1016/j.bbr.2018.11.045>
- Dorneich, M. C., Rogers, W., Whitlow, S. D. & DeMers, R. (2016). Human Performance Risks and Benefits of Adaptive Systems on the Flight Deck. *The International Journal of Aviation Psychology*, 26(1-2), 15–35. <https://doi.org/10.1080/10508414.2016.1226834>
- Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice / Andrew T. Duchowski*. Springer.
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological review*, 87(3), 272–300. <https://doi.org/10.1037/0033-295X.87.3.272>
- Durantin, G., Dehais, F. & Delorme, A. (2015). Characterization of mind wandering using fNIRS. *Frontiers in systems neuroscience*, 9, 45. <https://doi.org/10.3389/fnsys.2015.00045>
- Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D. & Manning, C. A. (1998). Situation Awareness as a Predictor of Performance for En Route Air Traffic Controllers. *Air Traffic Control Quarterly*, 6(1), 1–20. <https://doi.org/10.2514/atcq.6.1.1>
- Efron, B. & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.
- Egeth, H. & Kahneman, D. (1975). Attention and Effort. *The American Journal of Psychology*, 88(2), 339. <https://doi.org/10.2307/1421603>
- 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>
- Endsley, M. R. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 65–84. <https://doi.org/10.1518/001872095779049499>
- Endsley, M. R. (1995b). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64. <https://doi.org/10.1518/001872095779049543>
- Endsley, M. R. (2015). Situation Awareness Misconceptions and Misunderstandings. *Journal of Cognitive Engineering and Decision Making*, 9(1), 4–32. <https://doi.org/10.1177/1555343415572631>
- Endsley, M. R. (2019). A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. *Human Factors*, 18720819875376. <https://doi.org/10.1177/0018720819875376>
- Endsley, M. R. & Jones, D. G. (2012). *Designing for situation awareness: An approach to user-centered design / Mica R. Endsley and Debra G. Jones* (2nd ed.). CRC Press.
- Endsley, M. R. & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(2), 381–394. <https://doi.org/10.1518/001872095779064555>

- 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>
- F-35 Website. (2020, 23. März). <https://www.f35.com/about/capabilities/missionsystems>
- Feigh, K. M., Dorneich, M. C. & Hayes, C. C. (2012). Toward a characterization of adaptive systems: a framework for researchers and system designers. *Human Factors*, 54(6), 1008–1024. <https://doi.org/10.1177/0018720812443983>
- Ferrari, M. & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, 63(2), 921–935. <https://doi.org/10.1016/j.neuroimage.2012.03.049>
- Feuss, S. (2013). *Auf den ersten Blick: Wie Medieninhalte wahrgenommen und rezipiert werden*. Springer VS.
- Fortmann, F. & Lütke, A. (2013). An Intelligent SA-Adaptive Interface to Aid Supervisory Control of a UAV Swarm. *IEEE 11th International Conference on Industrial Informatics (INDIN)*, 768–773. <http://ieeexplore.ieee.org/servlet/opac?punumber=6599026>
- Fortmann, F. & Mengerlinghausen, T. (2014). Development and Evaluation of an Assistant System to Aid Monitoring Behavior during Multi-UAV Supervisory Control. In C. Stary (Hrsg.), *Proceedings of the 2014 European Conference on Cognitive Ergonomics - ECCE '14* (S. 1–8). ACM Press. <https://doi.org/10.1145/2637248.2637257>
- Fox, J., Merwin, D., Marsh, R., McConkie, G. & Kramer, A. (1996). Information Extraction during Instrument Flight: An Evaluation of the Validity of the Eye-Mind Hypothesis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 40(2), 77–81. <https://doi.org/10.1177/154193129604000215>
- Frische, F. & Lütke, A. (2013). SA-Tracer: A Tool for Assessment of UAV Swarm Operator SA during Mission Execution. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. <http://ieeexplore.ieee.org/servlet/opac?punumber=6520997>
- Gastwirth, J. L., Gel, Y. R. & Miao, W. (2009). The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. *Statistical Science*, 24(3). <https://doi.org/10.1214/09-STS301>
- Glaholt, M. G. (2014). *Eye tracking in the cockpit: a review of the relationships between eye movements and the aviators cognitive state*. <https://apps.dtic.mil/sti/citations/ad1000097>
- Goldberg, J. H. & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24(6), 631–645. [https://doi.org/10.1016/S0169-8141\(98\)00068-7](https://doi.org/10.1016/S0169-8141(98)00068-7)
- Grundgeiger, T., Hohm, A., Michalek, A., Egenolf, T., Markus, C. & Happel, O. (2021). The Validity of the SEEV Model as a Process Measure of Situation Awareness: The Example of a Simulated Endotracheal Intubation. *Human Factors*, 18720821991651. <https://doi.org/10.1177/0018720821991651>
- Haber, R. N. & Schindler, R. M. (1981). Error in proofreading: Evidence of syntactic control of letter processing? *Journal of Experimental Psychology: Human Perception and Performance*, 7(3), 573–579. <https://doi.org/10.1037/0096-1523.7.3.573>
- Haffaci, K., Massicotte, M.-C. & Doyon-Poulin, P. (2021). Goal-Directed Task Analysis for Situation Awareness Requirements During Ship Docking in Compulsory Pilotage Area. In (S. 647–654). Springer, Cham. [https://doi.org/10.1007/978-3-030-74608-7\\_79](https://doi.org/10.1007/978-3-030-74608-7_79)
- Hamid, H. & Waterson, P. (2011). Using Goal Directed Task Analysis to Identify Situation Awareness Requirements of Advanced Paramedics. In V. G. Duffy (Hrsg.), *Advances in Human Factors and Ergonomics Series. Advances in human factors and ergonomics in healthcare* (Bd. 20105254, S. 672–680). CRC. <https://doi.org/10.1201/EBK1439834978-c74>

- Hasanzadeh, S., Esmaili, B. & Dodd, M. D. (2016). Measuring Construction Workers' Real-Time Situation Awareness Using Mobile Eye-Tracking, 2894–2904. <https://doi.org/10.1061/9780784479827.288>
- Hauss, Y. & Eyferth, K. (2003). Securing future ATM-concepts' safety by measuring situation awareness in ATC. *Aerospace Science and Technology*, 7(6), 417–427. [https://doi.org/10.1016/S1270-9638\(02\)00011-1](https://doi.org/10.1016/S1270-9638(02)00011-1)
- Heilemann, F. (2022). *Skalierbare Delegation und adaptive Planungsassistenz zur Führung bemannt-unbemannter Teams: Konzept, Implementierung und Evaluation*.
- Helleberg, J. R. & Wickens, C. D. [Christopher D.] (2003). Effects of Data-Link Modality and Display Redundancy on Pilot Performance: An Attentional Perspective. *The International Journal of Aviation Psychology*, 13(3), 189–210. [https://doi.org/10.1207/S15327108IJAP1303\\_01](https://doi.org/10.1207/S15327108IJAP1303_01)
- Helton, W. S., Hollander, T. D., Warm, J. S., Tripp, L. D., Parsons, K., Matthews, G., Dember, W. N., Parasuraman, R [Raja] & Hancock, P. A [Peter A.] (2007). The abbreviated vigilance task and cerebral hemodynamics. *Journal of clinical and experimental neuropsychology*, 29(5), 545–552. <https://doi.org/10.1080/13803390600814757>
- Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F. & Schultz, T. (2013). Mental workload during n-back task-quantified in the prefrontal cortex using fNIRS. *Frontiers in human neuroscience*, 7, 935. <https://doi.org/10.3389/fnhum.2013.00935>
- Hoffelner, J. (2021). *Masterarbeit: Entwicklung und Evaluierung eines Messsystems zur robu. sten Abschätzung visueller Wahrnehmung im Cockpit*. Institut für Flugsysteme, Universität der Bundeswehr München.
- Hogg, D. N., Folles, K., Strand-Volden, F. & Torralba, B. (1995). Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics*, 38(11), 2394–2413.
- Honecker, F. & Schulte, A. (2017). Automated Online Determination of Pilot Activity Under Uncertainty by Using Evidential Reasoning. In D. Harris (Hrsg.), *Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics: Cognition and Design* (Bd. 10276, S. 231–250). Springer International Publishing. [https://doi.org/10.1007/978-3-319-58475-1\\_18](https://doi.org/10.1007/978-3-319-58475-1_18)
- Hornof, A. J. & Halverson, T. (2002). Cleaning up systematic error in eye-tracking data by using required fixation locations. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, 34(4), 592–604. <https://doi.org/10.3758/bf03195487>
- Huang, J [Jeff], White, R. & Buscher, G. (2012). User see, user point: gaze and cursor alignment in web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 1341–1350). <https://doi.org/10.1145/2207676.2208591>
- Huang, M. X., Kwok, T. C., Ngai, G., Chan, S. C. & Leong, H. V. (2016). Building a Personalized, Auto-Calibrating Eye Tracker from User Interactions. In *CHI '16, Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (S. 5169–5179). Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858404>
- Jackson, A. F. & Bolger, D. J. (2014). The neurophysiological bases of EEG and EEG measurement: a review for the rest of us. *Psychophysiology*, 51(11), 1061–1071. <https://doi.org/10.1111/psyp.12283>
- Johnson, A. W., Duda, K. R., Sheridan, T. B [Thomas B.] & Oman, C. M. (2017). A Closed-Loop Model of Operator Visual Attention, Situation Awareness, and Performance Across Automation Mode Transitions. *Human Factors*, 59(2), 229–241. <https://doi.org/10.1177/0018720816665759>
- Johnson, A. W., Oman, C. M., Sheridan, T. B [Thomas B.] & Duda, K. R. (2014). Dynamic task allocation in operational systems: Issues, gaps, and recommendations. *IEEE Conference on Aerospace*, 1–15. <https://doi.org/10.1109/AERO.2014.6836205>

- Jones, D. G. & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, space, and environmental medicine*, 67(6), 507–512.
- Jones, R. E. T., Connors, E. S., Mossey, M. E., Hyatt, J. R., Hansen, N. J. & Endsley, M. R. (2011). Using fuzzy cognitive mapping techniques to model situation awareness for army infantry platoon leaders. *Computational and Mathematical Organization Theory*, 17(3), 272–295. <https://doi.org/10.1007/s10588-011-9094-6>
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, 87(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kaber, D. B. & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, 5(2), 113–153. <https://doi.org/10.1080/1463922021000054335>
- Kharoufah, H., Murray, J., Baxter, G. & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences*, 99, 1–13. <https://doi.org/10.1016/j.paerosci.2018.03.002>
- Kidwell, B., Calhoun, G. L [Gloria L.], Ruff, H. A [Heath A.] & Parasuraman, R [Raja] (2012). Adaptable and Adaptive Automation for Supervisory Control of Multiple Autonomous Vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 428–432. <https://doi.org/10.1177/1071181312561096>
- Kim, H., Cheon, E.-J., Bai, D.-S., Lee, Y. H. & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry Investigation*, 15(3), 235–245. <https://doi.org/10.30773/pi.2017.08.17>
- Kim, J [Juno], Palmisano, S. A., Ash, A. & Allison, R. S. (2010). Pilot gaze and glideslope control. *ACM Transactions on Applied Perception*, 7(3), 1–18. <https://doi.org/10.1145/1773965.1773969>
- Kip, S. & Hancock, P. A [P. A.] (1995). Situation Awareness Is Adaptive, Externally Directed Consciousness. *Human Factors*, 37(1), 137–148.
- Kirwan, B. & Ainsworth, L. K. (1992). *A Guide To Task Analysis*. CRC Press. <https://doi.org/10.1201/b16826>
- Klaproth, O. W., Halbrügge, M., Krol, L. R., Vernaleken, C., Zander, T. O. & Russwinkel, N. (2020). A Neuroadaptive Cognitive Model for Dealing With Uncertainty in Tracing Pilots' Cognitive State. *Topics in cognitive science*, 12(3), 1012–1029. <https://doi.org/10.1111/tops.12515>
- Klein, G. (1993). A Recognition Primed Decision (RPD) Model of Rapid Decision Making.
- Klein, G., Phillips, J. K., Rall, E. L. & Peluso, D. A. (2007). A data-frame theory of sensemaking. In R. R. Hoffmann (Hrsg.), *Expertise Out Of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making* (S. 113–155).
- Knowles, W. B. (1963). OPERATOR LOADING TASKS. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 5, 155–161. <https://doi.org/10.1177/001872086300500206>
- Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A. & Kopacz, A. (2016). Discerning Ambient/Focal Attention with Coefficient K. *ACM Transactions on Applied Perception*, 13(3), 1–20. <https://doi.org/10.1145/2896452>
- Kunze, A., Summerskill, S. J., Marshall, R. & Filtness, A. J. (2019). Automation transparency: implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics*, 62(3), 345–360. <https://doi.org/10.1080/00140139.2018.1547842>
- Laberge, D. (1973). Attention and the measurement of perceptual learning. *Memory & cognition*, 1(3), 268–276. <https://doi.org/10.3758/BF03198108>
- Lavie, N., Hirst, A., Fockert, J. W. de & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of experimental psychology. General*, 133(3), 339–354. <https://doi.org/10.1037/0096-3445.133.3.339>

- Layton, C., Smith, P. J. & McCoy, C. E. (1994). Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 36(1), 94–119. <https://doi.org/10.1177/001872089403600106>
- Lindner, S. (2022). *Konzeption, Gestaltung und Evaluierung eines Human-Autonomy-Teaming-Systems für militärische Luftoperationen*.
- Lindner, S., Mund, D. & Schulte, A. (2022). How Human-Autonomy Teams change the Role of future Fighter Pilots: An Experimental Assessment. In *AIAA SCITECH Forum*. <https://doi.org/10.2514/6.2022-2551>
- Lounis, C., Hassoumi, A., Lefrancois, O., Peysakhovich, V. & Causse, M [Mickael] (2020). Detecting ambient/focal visual attention in professional airline pilots with a modified Coefficient K: a full flight simulator study. In A. Bulling, A. Huckauf, E. Jain, R. Radach & D. Weiskopf (Hrsg.), *Symposium on Eye Tracking Research and Applications* (S. 1–6). ACM. <https://doi.org/10.1145/3379157.3391412>
- Lounis, C., Peysakhovich, V. & Causse, M [Mickaël]. (2020a). Flight Eye Tracking Assistant (FETA): Proof of Concept. In N. A. Stanton (Hrsg.), *Advances in Intelligent Systems and Computing. Advances in Human Factors of Transportation* (Bd. 964, S. 739–751). Springer International Publishing. [https://doi.org/10.1007/978-3-030-20503-4\\_66](https://doi.org/10.1007/978-3-030-20503-4_66)
- Lounis, C., Peysakhovich, V. & Causse, M [Mickaël]. (2020b). *Lempel-Ziv Complexity of dwell sequences: visual scanning pattern differences between novice and expert aircraft pilots*. <https://doi.org/10.3929/ETHZ-B-000407653>
- Lounis, C., Peysakhovich, V. & Causse, M [Mickaël] (2021). Visual scanning strategies in the cockpit are modulated by pilots' expertise: A flight simulator study. *PloS one*, 16(2), e0247061. <https://doi.org/10.1371/journal.pone.0247061>
- Mack, A. (2003). Inattentive Blindness. *Current Directions in Psychological Science*, 12(5), 180–184. <https://doi.org/10.1111/1467-8721.01256>
- Maiwald, F. & Schulte, A. (2012). Adaptation of a Human Resource Model by the Use of Machine Learning Methods as Part of a Military Helicopter Pilot Associate System. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 970–974. <https://doi.org/10.1177/1071181312561203>
- Manzey, D., Reichenbach, J. & Onnasch, L. (2012). Human Performance Consequences of Automated Decision Aids. *Journal of Cognitive Engineering and Decision Making*, 6(1), 57–87. <https://doi.org/10.1177/1555343411433844>
- Martel, A., Dähne, S. & Blankertz, B. (2014). EEG predictors of covert vigilant attention. *Journal of Neural Engineering*, 11(3), 35009. <https://doi.org/10.1088/1741-2560/11/3/035009>
- Martinez-Conde, S. & Macknik, S. L. (2015). From Exploration to Fixation: An Integrative View of Yarbus's Vision. *Perception*, 44(8-9), 884–899. <https://doi.org/10.1177/0301006615594963>
- Martinez-Conde, S., Macknik, S. L., Troncoso, X. G. & Dyar, T. A. (2006). Microsaccades Counteract Visual Fading during Fixation. *Neuron*, 49(6), 929. <https://doi.org/10.1016/j.neuron.2006.02.007>
- Matias, J., Belletier, C., Izaute, M., Lutz, M. & Silvert, L. (2021). The role of perceptual and cognitive load on inattentive blindness: A systematic review and three meta-analyses. *Quarterly journal of experimental psychology* (2006), 17470218211064903. <https://doi.org/10.1177/17470218211064903>
- Matthews, M. D. & Beal, S. A. (2002). *Assessing Situation Awareness in Field Training Exercises*. Defense Technical Information Center. <https://doi.org/10.21236/ada408560>
- McVay, J. C. & Kane, M. J. (2012). Drifting from slow to "D'oh!": working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of experimental*

- psychology. Learning, memory, and cognition*, 38(3), 525–549.  
<https://doi.org/10.1037/a0025896>
- Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8(4), 283–298.  
[https://doi.org/10.1016/s0001-2998\(78\)80014-2](https://doi.org/10.1016/s0001-2998(78)80014-2)
- Molloy, R. & Parasuraman, R [Raja] (1996). Monitoring an Automated System for a Single Failure: Vigilance and Task Complexity Effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38(2), 311–322. <https://doi.org/10.1518/001872096779048093>
- Moore, K. & Gugerty, L. (2010). Development of a novel measure of situation awareness: The case for eye movement analysis. *Proceedings of the Human Factors and Ergonomics Society*, 54.
- Moray, N. (1959). Attention in Dichotic Listening: Affective Cues and the Influence of Instructions. *Quarterly Journal of Experimental Psychology*, 11(1), 56–60.  
<https://doi.org/10.1080/17470215908416289>
- Moray, N. (1970). Towards a quantitative theory of attention. *Acta Psychologica*, 33, 111–117.  
[https://doi.org/10.1016/0001-6918\(70\)90126-5](https://doi.org/10.1016/0001-6918(70)90126-5)
- Mund, D. (2023). *Bewertung psychophysiologischer Mess-Sensoren als Basis für nutzeradaptive Automation*.
- Murphy, G., Groeger, J. A. & Greene, C. M. (2016). Twenty years of load theory-Where are we now, and where should we go next? *Psychonomic bulletin & review*, 23(5), 1316–1340.  
<https://doi.org/10.3758/s13423-015-0982-5>
- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology* / Ulric Neisser. W. H. Freeman.
- Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, 2(1), Artikel 36. <https://doi.org/10.5334/joc.58>
- O'Neill, T., McNeese, N., Barron, A. & Schelble, B. (2022). Human-Autonomy Teaming: A Review and Analysis of the Empirical Literature. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 64(5), 904–938. <https://doi.org/10.1177/0018720820960865>
- Onken, R. & Schulte, A. (2010). *System-ergonomic design of cognitive automation: Dual-mode cognitive design of vehicle guidance and control work systems* / Reiner Onken and Axel Schulte. *Studies in computational intelligence*, 1860-949X: v. 235. Springer. <http://www.springer.com/gb/BLDSS>
- Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J [Jeff] & Hays, J. (2016). WebGazer: Scalable Webcam Eye Tracking Using User Interactions. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence* (S. 3839–3845).
- Parasuraman, R [Raja], Cosenzo, K. A. & Visser, E. de (2009). Adaptive Automation for Human Supervision of Multiple Uninhabited Vehicles: Effects on Change Detection, Situation Awareness, and Mental Workload. *Military Psychology*, 21(2), 270–297.  
<https://doi.org/10.1080/08995600902768800>
- Parasuraman, R [Raja] & Manzey, D. H. (2010). Complacency and bias in human use of automation: an attentional integration. *Human Factors*, 52(3), 381–410.  
<https://doi.org/10.1177/0018720810376055>
- Parasuraman, R [Raja], Molloy, R. & Singh, I. L. (1993). Performance Consequences of Automation-Induced 'Complacency'. *The International Journal of Aviation Psychology*, 3(1), 1–23.  
[https://doi.org/10.1207/s15327108ijap0301\\_1](https://doi.org/10.1207/s15327108ijap0301_1)
- Parasuraman, R [Raja], Sheridan, T. B [T. B.] & Wickens, C. D [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>
- Pashler, H. E. (1998). *The psychology of attention. A Bradford book*. MIT.

- Patel, M., Lal, S., Kavanagh, D. & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*, 38(6), 7235–7242. <https://doi.org/10.1016/j.eswa.2010.12.028>
- Peißl, S., Wickens, C. D [Christopher D.] & Baruah, R. (2018). Eye-Tracking Measures in Aviation: A Selective Literature Review. *The International Journal of Aerospace Psychology*, 28(3-4), 98–112. <https://doi.org/10.1080/24721840.2018.1514978>
- R.M. Taylor. (2017). Situational Awareness Rating Technique (Sart): The Development of a Tool for Aircrew Systems Design. In *Situational Awareness* (S. 111–128). Routledge. <https://doi.org/10.4324/9781315087924-8>
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3), 257–266. <https://doi.org/10.1109/TSMC.1983.6313160>
- Ratwani, R. M., McCurry, J. M. & Trafton, J. G. (2010). Single Operator, Multiple Robots: An Eye Movement Based Theoretic Model of Operator Situation Awareness. *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. <http://ieeexplore.ieee.org/servlet/opac?punumber=5447157>
- Reason, J. (1990a). The contribution of latent human failures to the breakdown of complex systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 327(1241), 475–484. <https://doi.org/10.1098/rstb.1990.0090>
- Reason, J. (1990b). *Human error*. Cambridge University Press.
- Regis, N., Dehais, F., Rachelson, E., Thooris, C., Pizziol, S., Causse, M [Mickael] & Tessier Catherine (2014). Formal Detection of Attentional Tunneling in Human Operator–Automation Interactions. *IEEE Transactions on Human-Machine Systems*, 44(3), 326–336. <https://doi.org/10.1109/THMS.2014.2307258>
- Remington, R. W. & Loft, S. (2015). Attention and multitasking. In D. A. Boehm-Davis, F. T. Durso & J. D. Lee (Hrsg.), *APA handbook of human systems integration* (S. 261–276). American Psychological Association. <https://doi.org/10.1037/14528-017>
- Richard E. Clark, David F. Feldon, Jeroen J. G. van Merriënboer, Kenneth A. Yates & Sean Early. (2001). Cognitive Task Analysis. In M. P. Driscoll, D. H. Jonassen & P. Harris (Hrsg.), *AECT Series, 2. Handbook of Research for Educational Communications and Technology: A Project of the Association for Educational Communications and Technology* (3. Aufl., S. 577–593). Taylor & Francis. <https://doi.org/10.4324/9780203880869-48>
- Rouse, W. B. (1988). Adaptive Aiding for Human/Computer Control. *Human Factors*.
- Salas, E [Eduardo]. (2017). *Situation Awareness in Team Performance: Implications for Measurement and Training* (First edition). Routledge.
- Salmon, P. M [Paul M.], Stanton, N. A., Walker, G. & Green, D. (2006). Situation awareness measurement: a review of applicability for C4i environments. *Applied ergonomics*, 37(2), 225–238. <https://doi.org/10.1016/j.apergo.2005.02.001>
- Salmon, P. M [Paul M.], Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L. & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, 39(3), 490–500. <https://doi.org/10.1016/j.ergon.2008.10.010>
- Sarter, N. B., Mumaw, R. J. & Wickens, C. D [Christopher D.] (2007). Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data. *Human Factors*, 49(3), 347–357. <https://doi.org/10.1518/001872007X196685>
- Sarter, N. B., Schroeder, B. & McGuirl, J. (2001). Supporting decision-making and action selection under time pressure and uncertainty - The case of inflight icing. In *39th Aerospace Sciences Meeting*

- and Exhibit*. American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/6.2001-543>
- Sarter, N. B. & Woods, D. D. (1991). Situation Awareness: A Critical But Ill-Defined Phenomenon. *The International Journal of Aviation Psychology*, 1, 45–57.
- Sauer, J., Chavaillaz, A. & Wastell, D. (2017). On the effectiveness of performance-based adaptive automation. *Theoretical Issues in Ergonomics Science*, 18(3), 279–297. <https://doi.org/10.1080/1463922X.2016.1207824>
- Sauvet, F., Bougard, C., Coroenne, M., Lely, L., van Beers, P., Elbaz, M., Guillard, M., Leger, D. & Chennaoui, M. (2014). In-flight automatic detection of vigilance states using a single EEG channel. *IEEE transactions on bio-medical engineering*, 61(12), 2840–2847. <https://doi.org/10.1109/TBME.2014.2331189>
- Schardt, W. A., Caufield, K. J. & Dyre, B. P. (2002). Effects of a Virtual Air Speed Error Indicator on Guidance Accuracy and Eye Movement Control during Simulated Flight. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46(17), 1594–1598. <https://doi.org/10.1177/154193120204601714>
- Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, 84(1), 1–66. <https://doi.org/10.1037/0033-295x.84.1.1>
- Schulte, A. (2013). *Kognitive und kooperative Automation zur Führung unbemannter Luftfahrzeuge*. <https://doi.org/10.17185/dupublico/31350>
- Schulte, A., Donath, D. & Lange, D. S. (2016). Design Patterns for Human-Cognitive Agent Teaming. In D. Harris (Hrsg.), *Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics* (Bd. 9736, S. 231–243). Springer International Publishing. [https://doi.org/10.1007/978-3-319-40030-3\\_24](https://doi.org/10.1007/978-3-319-40030-3_24)
- Schwarz, J. & Fuchs, S. (2017). Multidimensional Real-Time Assessment of User State and Performance to Trigger Dynamic System Adaptation. In D. D. Schmorow & C. M. Fidopiastis (Hrsg.), *Lecture Notes in Computer Science. Augmented Cognition. Neurocognition and Machine Learning* (Bd. 10284, S. 383–398). Springer International Publishing. [https://doi.org/10.1007/978-3-319-58628-1\\_30](https://doi.org/10.1007/978-3-319-58628-1_30)
- Schwerd, S. & Schulte, A. (2020). Experimental Validation of an Eye-Tracking-Based Computational Method for Continuous Situation Awareness Assessment in an Aircraft Cockpit. In D. Harris & W.-C. Li (Hrsg.), *Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics. Cognition and Design* (Bd. 12187, S. 412–425). Springer International Publishing. [https://doi.org/10.1007/978-3-030-49183-3\\_32](https://doi.org/10.1007/978-3-030-49183-3_32)
- Schwerd, S. & Schulte, A. (2021a). Measuring the Deviation between Ground Truth and Operator Awareness in a UAV Management Scenario: An Eye-Tracking Approach. In *AIAA Scitech 2021 Forum*. American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/6.2021-1579>
- Schwerd, S. & Schulte, A. (2021b). Operator State Estimation to Enable Adaptive Assistance in Manned-Unmanned-Teaming. *Cognitive Systems Research*, 67, 73–83. <https://doi.org/10.1016/j.cogsys.2021.01.002>
- Schwerd, S. & Schulte, A. (2021c). Situation Awareness Adaptive Alerting in an Aircraft Cockpit: A Simulator Study. *Preprint*. Vorab-Onlinepublikation. <https://doi.org/10.31224/osf.io/9gmnv>
- Schwerd, S. & Schulte, A. (2022). Experimental Assessment of Fixation-Based Attention Measurement in an Aircraft Cockpit. In D. Harris & W.-C. Li (Hrsg.), *Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics* (Bd. 13307, S. 408–419). Springer International Publishing. [https://doi.org/10.1007/978-3-031-06086-1\\_32](https://doi.org/10.1007/978-3-031-06086-1_32)

- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shappell, S. A. & Wiegmann, D. A. (1998). *A Human Error Analysis of General Aviation Controlled Flight into Terrain Accidents Occurring Between 1990-1998* (DOT/FAA/AM-03/4).
- Sheridan, T. B [Thomas B.]. (1992). *Telerobotics, automation, and human supervisory control*. MIT Press.
- Shiferaw, B., Downey, L. & Crewther, D. (2019). A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience and biobehavioral reviews*, 96, 353–366. <https://doi.org/10.1016/j.neubiorev.2018.12.007>
- Shorrock, S. T. (2007). Errors of perception in air traffic control. *Safety Science*, 45(8), 890–904. <https://doi.org/10.1016/j.ssci.2006.08.018>
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489. <https://doi.org/10.1038/nature16961>
- Simons, D. J. & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentive blindness for dynamic events. *Perception*, 28(9), 1059–1074. <https://doi.org/10.1068/p2952>
- Spearman, C. (1961). The Proof and Measurement of Association Between Two Things. In *Studies in individual differences: The search for intelligence* (S. 45–58). Appleton-Century-Crofts. <https://doi.org/10.1037/11491-005>
- Stanton, N. A. (2006). Hierarchical task analysis: developments, applications, and extensions. *Applied ergonomics*, 37(1), 55–79. <https://doi.org/10.1016/j.apergo.2005.06.003>
- Stanton, N. A. (2016). Distributed situation awareness. *Theoretical Issues in Ergonomics Science*, 17(1), 1–7. <https://doi.org/10.1080/1463922X.2015.1106615>
- Stanton, N. A., Salmon, P. M [P. M.], Walker, G. H [G. H.], Salas, E [E.] & Hancock, P. A [P. A.] (2017). State-of-science: situation awareness in individuals, teams and systems. *Ergonomics*, 60(4), 449–466. <https://doi.org/10.1080/00140139.2017.1278796>
- Steelman-Allen, K. S. & McCarley, J. S. (2011). Modeling the control of attention in complex visual displays. *Human Factors*, 53, 143–153.
- Strasburger, H., Rentschler, I. & Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *Journal of vision*, 11(5), 13. <https://doi.org/10.1167/11.5.13>
- Styles, E. A. (2005). *The psychology of attention* (2nd ed.). Psychology. <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10152267>
- Sugano, Y., Matsushita, Y., Sato, Y. & Koike, H. (2008). An Incremental Learning Method for Unconstrained Gaze Estimation. *European Conference on Computer Vision*, 5304, 656–667. [https://doi.org/10.1007/978-3-540-88690-7\\_49](https://doi.org/10.1007/978-3-540-88690-7_49)
- Tanner, W. P. & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological review*, 61(6), 401–409. <https://doi.org/10.1037/h0058700>
- Thomson, D. R., Besner, D. & Smilek, D. (2015). A resource-control account of sustained attention: evidence from mind-wandering and vigilance paradigms. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 10(1), 82–96. <https://doi.org/10.1177/1745691614556681>
- Tjerk, E. G. de, Henryk, F. R. A. & Neerincx, M. A. (2010). Adaptive Automation Based on an Object-Oriented Task Model: Implementation and Evaluation in a Realistic C2 Environment. *Journal of Cognitive Engineering and Decision Making*, 4(2), 152–182. <https://doi.org/10.1518/155534310X522851>

- Tulving, E., Mandler, G. & Bauml, R. (1964). Interaction of Two Sources of Information in Tachistoscopic Word Recognition. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 18(1), 62–71. <https://doi.org/10.1037/h0083296>
- Underwood, G. & Everatt, J. (1992). The Role of Eye Movements in Reading: Some Limitations of the Eye-Mind Assumption. In E. Chekaluk & K. Llewellyn (Hrsg.), *Advances in Psychology: Bd. 88. The Role of eye movements in perceptual processes* (Bd. 88, S. 111–169). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61744-6](https://doi.org/10.1016/S0166-4115(08)61744-6)
- van de Merwe, K., van Dijk, H. & Zon, R. (2012). Eye Movements as an Indicator of Situation Awareness in a Flight Simulator Experiment. *The International Journal of Aviation Psychology*, 22(1), 78–95. <https://doi.org/10.1080/10508414.2012.635129>
- Verdière, K. J., Roy, R. N [Raphaëlle N.] & Dehais, F. (2018). Detecting Pilot's Engagement Using fNIRS Connectivity Features in an Automated vs. Manual Landing Scenario. *Frontiers in human neuroscience*, 12, 6. <https://doi.org/10.3389/fnhum.2018.00006>
- Warm, J. S., Parasuraman, R [Raja] & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, 50(3), 433–441. <https://doi.org/10.1518/001872008X312152>
- Wei, H., Zhuang, D., Wanyan, X. & Wang, Q. (2013). An experimental analysis of situation awareness for cockpit display interface evaluation based on flight simulation. *Chinese Journal of Aeronautics*, 26(4), 884–889. <https://doi.org/10.1016/j.cja.2013.04.053>
- Welford, A. T. (1952). THE 'PSYCHOLOGICAL REFRACTORY PERIOD' AND THE TIMING OF HIGH-SPEED PERFORMANCE-A REVIEW AND A THEORY. *British Journal of Psychology. General Section*, 43(1), 2–19. <https://doi.org/10.1111/j.2044-8295.1952.tb00322.x>
- Wickens, C. D [Christopher D.] (2005). Attentional Tunneling and Task Management. *International Symposium on Aviation Psychology*.
- Wickens, C. D [Christopher D.]. (2015). Noticing Events in the Visual Workplace: The SEEV and NSEEV Models. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman & J. L. Szalma (Hrsg.), *The Cambridge Handbook of Applied Perception Research* (S. 749–768). Cambridge University Press. <https://doi.org/10.1017/CBO9780511973017.046>
- Wickens, C. D [Christopher D.] (2021). Attention: Theory, Principles, Models and Applications. *International Journal of Human-Computer Interaction*, 37(5), 403–417. <https://doi.org/10.1080/10447318.2021.1874741>
- Wickens, C. D [Christopher D.], Alexander, A. L., Horrey, W. J., Nunes, A. & Hardy, T. J. (2004). Traffic and Flight Guidance Depiction on a Synthetic Vision System Display: The Effects of Clutter on Performance and Visual Attention Allocation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(1), 218–222. <https://doi.org/10.1177/154193120404800147>
- Wickens, C. D [Christopher D.], Clegg, B. A., Vieane, A. Z. & Sebok, A. L. (2015). Complacency and Automation Bias in the Use of Imperfect Automation. *Human Factors*, 57(5), 728–739. <https://doi.org/10.1177/0018720815581940>
- Wickens, C. D [Christopher D.], Hollands, J. G., Banbury, S. & Parasuraman, R [R.]. (2022). *Engineering psychology and human performance* (Fifth Edition). Pearson.
- Wickens, C. D [Christopher D.], McCarley, J. S., Alexander, A. L., Thomas, Lisa C, Ambinder, Michael & Zheng, S. (2008). Attention-Situation Awareness Model of Pilot Error. *Human performance modeling in aviation*, 213–239.
- Wickens, C. D [Christopher D.], Sebok, A., Bagnall, T. & Kamienski, J. (2007). Modeling of Situation Awareness Supported by Advanced Flight Deck Displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 51(12), 794–798. <https://doi.org/10.1177/154193120705101204>

- Wiegmann, D. A. & Shappell, S. A. (2003). A Human Error Approach to Aviation Accident Analysis. Vorab-Onlinepublikation. <https://doi.org/10.4324/9781315263878>
- Wiener, E. L. (1981). *Complacency: Is the term useful for air safety*.
- William James. (1890). *The Principles of Psychology*.
- Wilson, G. F. & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Human Factors*, 49(6), 1005–1018. <https://doi.org/10.1518/001872007X249875>
- Winter, J. C. F. de, Eisma, Y. B., Cabrall, C. D. D., Hancock, P. A [P. A.] & Stanton, N. A. (2019). Situation awareness based on eye movements in relation to the task environment. *Cognition, Technology & Work*, 21(1), 99–111. <https://doi.org/10.1007/s10111-018-0527-6>
- Wolfe, J. M. & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3). <https://doi.org/10.1038/s41562-017-0058>
- Yantis, S. & Johnston, J. C. (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 16(1), 135–149. <https://doi.org/10.1037/0096-1523.16.1.135>
- Zhang, T., Yang, J., Liang, N., Pitts, B. J., Prakah-Asante, K. O., Curry, R., Duerstock, B. S., Wachs, J. P. & Yu, D. (2020). Physiological Measurements of Situation Awareness: A Systematic Review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 18720820969071. <https://doi.org/10.1177/0018720820969071>
- Zhang, X., Huang, M. X., Sugano, Y. & Bulling, A. (2018). Training Person-Specific Gaze Estimators from User Interactions with Multiple Devices. In R. Mandryk, M. Hancock, M. Perry & A. Cox (Hrsg.), *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (S. 1–12). ACM. <https://doi.org/10.1145/3173574.3174198>
- Ziv, G. (2016). Gaze Behavior and Visual Attention: A Review of Eye Tracking Studies in Aviation. *The International Journal of Aviation Psychology*, 26(3-4), 75–104. <https://doi.org/10.1080/10508414.2017.1313096>

## C. ABBILDUNGSVERZEICHNIS

Abbildung 2-1 - Ein Model der menschlichen Informationsverarbeitung nach (Remington & Loft, 2015; Wickens et al., 2022).....	4
Abbildung 2-2 - Filtertheorie nach Broadbent (Styles, 2005, S. 21) .....	6
Abbildung 2-3 - Überwachende Kontrolle eines technischen Prozesses nach Sheridan (1992) .....	8
Abbildung 2-4 - Schematische Abbildung von Fixationsfolgen innerhalb und zwischen Interessensgebieten, welche durch Sakkaden verbunden sind.....	13
Abbildung 2-5 - Vergleich der Gestaltung zweier unterschiedlicher Fahrtmesser in einem HUD (Schaudt et al., 2002): Konventionell (links) gegenüber der neuartigen Gestaltung (rechts) .....	15
Abbildung 2-6 - Prozess und Einflussfaktoren bei der Entscheidungsfindung eines Bedieners (Endsley & Jones, 2012).....	19
Abbildung 2-7 - Der Wahrnehmungszyklus nach Adams et al. (1995) .....	20
Abbildung 2-8 - Risikoraum eines Fluglotsen bei der Überwachung der räumlichen Trennung zweier Flugzeuge (Kip & Hancock, 1995).....	21
Abbildung 2-9 - Attention-SA Modell nach Wickens et al. (2008) .....	22
Abbildung 2-10 - Konzeptionelle <i>FCM Submap</i> aus der Arbeit von R. E. T. Jones et al. (2011).....	23
Abbildung 2-11 - Die wachsende Unsicherheit über einen Systemzustand zwischen Fixationen (Johnson et al., 2017).....	24
Abbildung 2-12 - Hypothetische Abbildung einer Situationsbewusstseinsmessung basierend auf der SAGAT (Winter et al., 2019) .....	25
Abbildung 2-13 - Systemaufbau des techn. Messsystems aus Fortmann und Mengerlinghausen (2014) .....	28
Abbildung 2-14 - Generischer Aufbau eines adaptiven Systems adaptiert von Feigh et al. (2012) .....	37
Abbildung 3-1 - Struktur des adaptiven Systems bestehend aus Aufmerksamkeitsanalyse, SB-Bewertung und Intervention, die sich auf die Mensch-Maschine-Schnittstelle (MMS) auswirkt. ....	42
Abbildung 3-2 - Messkette zur kontinuierlichen Analyse von visueller Aufmerksamkeit. ....	43
Abbildung 3-3 - Vergleich einer Fixation mit zwei Auswertungsradien. Beim Radius 1 liegt die Aol 1 nicht innerhalb des Auswertungsradius und wird somit nicht dieser Fixation zugewiesen. Bei Radius 2 liegt die Aol 1 innerhalb des Radius und wird somit als Fixation auf Aol 1 verarbeitet. ....	44
Abbildung 3-4 - Beispiel der semantischen Auswertung zweier hypothetischer Fixationen im Cockpit eines Autos.....	45
Abbildung 3-5 - Messkette zur Bewertung von Situationsbewusstsein. ....	46
Abbildung 3-6 - Ein Blick auf die taktische Karte mit zwei UAVs und der resultierenden Informationen. Abbildung adaptiert aus Schwerd und Schulte (2021a) .....	48
Abbildung 3-7 - Grundlegende Bestandteile des statischen Modells.....	48
Abbildung 3-8 - Beispiel eines SB-Modells mit Knoten und den assoziierten Observationssemantiken. ....	50
Abbildung 3-9 - Beispiel einer Inferenz basierend auf dem Modell in Abbildung 3-8.....	51
Abbildung 3-10 - Berechnung einer Abweichung im Knoten „Höhentrennung“ zwischen den Observation 1 und 3 mit exemplarischen Zeitverlauf der Abweichungsberechnung. ....	52
Abbildung 3-11 - Adaptionsmechanismus mit den Eingangsgrößen SB-Bewertung sowie der Analyse von Missionskontext und Situation. Die wirkt auf die MMS. ....	54
Abbildung 3-12 - Gesamtsystemarchitektur des SB-adaptiven Assistenzsystems. ....	55
Abbildung 3-13 - Pseudoklassendiagramm einer <i>Observation</i> (links) und einer <i>SitAwNode</i> (rechts).....	56
Abbildung 3-14 - Klassendiagramm der <i>SitAwMap</i> (links). Beispiel wie mithilfe der Transitionsmatrizen ein Netz modelliert werden kann (rechts).....	57

Abbildung 3-15 - Klassendiagramm der Komponenten zur dynamischen Berechnung der SB-Modelle .....	57
Abbildung 4-1 - Der Cockpitsimulator am IFS besteht aus einem Sichtsystem mit HUD, drei Touchscreens sowie einem HOTAS System und Ruderpedalen zur Bedienung. ....	61
Abbildung 4-2 - Weltmodell des Blickbewegungsmesssystems inklusive Positionierung der Infrarotstrahler und Kameras. ....	62
Abbildung 4-3 - Systemarchitektur der Simulationsumgebung bestehend aus Simulations- und Automationsapplikationen (grün), der MMS (lila) sowie der Applikationen für die adaptive Assistenz (orange).....	63
Abbildung 4-4 - Probekbasiertes Ansatz zur Auswertung des Radius um eine gemessene Fixation. Proben werden an den Schnittpunkten der grauen Linien durchgeführt. ....	64
Abbildung 4-5 - Oberfläche des SitAwEditors.....	65
Abbildung 4-6 - Versuchskonzept zur objektiven Messung von Aufmerksamkeit im natürlichen Arbeitsumgebung.....	67
Abbildung 4-7 - Abzufliegende Strecke der Mission 1 bestehend aus zehn Wegpunkten .....	68
Abbildung 4-8 - Übersicht über die Lage in der Mission 2.....	69
Abbildung 4-9 - Gestaltung der Oberfläche für Mission 1 und 2. Die blau markierten Interessensgebiete sind Fluggeschwindigkeit (VEL, 1) Flughöhe (ALT, 2), Nachricht (MSG, 3), Flugpfad (ROUTE, 4) und taktisches Objekt (TO, 5).....	70
Abbildung 4-10 - Verlauf von Flughöhe, Steuerknüppelausschlag, Schubhebelstellung und Flugzeugneigung nach einer Windböe bei T=1064s mit markierter Reaktion in der Steuerknüppelstellung ist markiert. ....	71
Abbildung 4-11 - Zeit $T_{fix} - react$ zwischen Beginn der ersten Fixation und Reaktion in Abhängigkeit von Aufgabe und Auswertungsradius. Referenz auf null (rot-gestrichelt) und Referenz auf das obere Quartil der Reaktionszeit $T_{task}(AoI)$ (grün-gestrichelt). ....	74
Abbildung 4-12 - Oben: Verhältnis von Fehler- zu Trefferrate in Abhängigkeit vom Auswertungsradius. Unten: Präzision und Sensitivität im Verhältnis zum Auswertungsradius.....	75
Abbildung 4-13 - Vergleich aller Fixationsdauern mit den Fixationsdauern vor Reaktion auf die Aufgaben ROUTE, ALT, TO, VEL und MESSAGE.....	76
Abbildung 4-14 - Winkeldistanz des Blicks zum aufgabenbezogenen Interessensgebiet zum Zeitpunkt einer Aktion.....	77
Abbildung 4-15 - Flugpfad definiert durch blaue Wegpunkte und Platzierung der auftauchenden Flugabwehrstellungen in Missionsszenario 1 .....	81
Abbildung 4-16 - Platzierung der UAVs und der auftauchenden Ziele zur Aufklärung im Missionsszenario 2.....	82
Abbildung 4-17 - SB-Modelle für Mission 1 (links) und Mission 2 (rechts). Die Normierungskonstante ist bei jedem SB-Knoten unten links annotiert (lila). ....	83
Abbildung 4-18 - Zeitstrahl der Aufgabe in Mission 2 mit gelb markierten Delegations- und Klassifikationszeit.....	84
Abbildung 4-19 - Individuelle Genauigkeit der Blickbewegungsmessung quantifiziert durch den durchschnittlichen Fehler und Standardabweichung.....	85
Abbildung 4-20 - Zusammenhang zwischen mittlerer Abweichung im Knoten Radarhöhe und der relativen Fehlerzeit sowie dem relativen SART-Ergebnis. ....	87
Abbildung 4-21 - Klassifikationsleistung und Positivrate über den individuellen Teilnehmern (P=Pilot, S=Student).....	88
Abbildung 4-22 - Zeitdifferenz zwischen Meldezeitpunkt und dem Zeitpunkt der ersten <i>devBedrohung</i> = 0 im Knoten „Bedrohung“ der zugehörigen SAM-Site-Identifikation (n=157) .....	89

Abbildung 4-23 - Abweichung und Radarhöhe zum Zeitpunkt ausgewählter Fehlerfälle in Mission 1 für die Versuchspersonen P1-7 und S1. Bei P2 liegt ein Messfehler vor. Es gilt zu beachten, dass die individuelle y-Achsen in der Radarhöhe unterschiedlich sind, da die Fehlerhöhen unterschiedlich groß sind.....	90
Abbildung 4-24 - Abweichung und Radarhöhe zum Zeitpunkt ausgewählter Fehlerfälle in Mission 1 für die Versuchspersonen S2-S9.....	91
Abbildung 4-25 - Zeitdifferenz zwischen Auftauchen einer Bedrohung der Meldung. In orange markiert ist der Grenzwert, ab dem eine Zeitdifferenz als Fehler klassifiziert wird. ....	92
Abbildung 4-26 - Abweichungswerte von P1 in den Knoten „Missionszeit“, „Position“, „Bedrohung#2“ und „Radarhöhe“ für den Zeitraum ab Auftauchen der Bedrohung#2 (orange, gestrichelt) bis kurz nach Meldezeitpunkt (blau, gestrichelt).....	93
Abbildung 4-27 - Abweichungswerte von P1 in den Knoten „Missionszeit“, „Position“, „Bedrohung#1“ und „Radarhöhe“ für den Zeitraum ab Auftauchen der Bedrohung#1 (orange, gestrichelt) bis kurz nach Meldezeitpunkt (blau, gestrichelt).....	94
Abbildung 4-28 - Zusammenhang zwischen mittlerer Delegationszeit und mittlerer SB-Abweichung im Knoten „UAV-Aktivität“.....	95
Abbildung 4-29 - Zusammenhang zwischen relativem SART-Ergebnis und mittlerer SB-Abweichung im Knoten „UAV Aktivität“.....	96
Abbildung 4-30 - Zeitdifferenz zwischen Auftauchen und Auswahl eines Ziels. In orange markiert ist der Grenzwert, ab dem eine Zeitdifferenz als Fehler klassifiziert wird ( $T > 20$ )......	97
Abbildung 4-31 - Zeitdifferenz zwischen dem Zeitpunkt der gemessenen Wahrnehmung im SB-Modell und der Auswahl eines Ziels für alle Datenpunkte mit Fehler.....	97
Abbildung 4-32 - Abweichungswerte in den Knoten „UAV Aktivität“ und „UAV Position“ für korrekte und inkorrekte Entscheidungen.....	98
Abbildung 4-33 - Zeitdifferenz zwischen Auftauchen und Auswahl eines Aufklärungsbilds. In orange markiert ist der Grenzwert, ab dem eine Zeitdifferenz als Fehler klassifiziert wird ( $T > 40$ ). ....	98
Abbildung 4-34 - Zeitdifferenz zwischen dem Zeitpunkt der gemessenen Wahrnehmung im SB-Modell und dem Beginn der Klassifikation für alle Datenpunkte mit Fehler.....	99
Abbildung 4-35 - Wegpunkte in der Mission 1 .....	103
Abbildung 4-36 - Übersicht über die Lage in der Mission 2.....	103
Abbildung 4-37 - SB Modell des ersten Versuchs. Der Parameter $cn$ ist in lila markiert. ....	104
Abbildung 4-38 - Darstellung der Flugpfadabweichung .....	104
Abbildung 4-39 - SB Modell des zweiten Versuchs.....	104
Abbildung 4-40 - Hervorhebung von taktischen Objekten: Gebäude (links) und Flugabwehrstellung (rechts).....	105
Abbildung 4-41 - Pfeil am Rand der taktischen Karte im MHDD, der die Richtung des hervorgehobenen Objekts visualisiert.....	105
Abbildung 4-42 - Zusammenhang zwischen mittleren Fehlern in den drei Verfolgungsaufgaben Flughöhe, Geschwindigkeit sowie Flugpfad und der SB-Abweichung in der damit verbundenen, relevanten Systeminformation .....	108
Abbildung 4-43 - Zusammenhang zwischen Anzahl an Interventionen und der SB-Abweichung in den drei Verfolgungsaufgaben Flughöhe, Geschwindigkeit und Flugpfad .....	108
Abbildung 4-44 - Vergleich der Leistung in den drei Verfolgungsaufgaben zwischen der Mission mit und ohne SB-adaptive Intervention. Die Kreuze stehen für einen Probanden und sind über die zwei Missionen miteinander verbunden, um die individuelle Veränderung der Leistung pro Probanden zu visualisieren. ....	109
Abbildung 4-45 - Erkennungszeiten in den drei Aufgaben Flughöhe, Geschwindigkeit und Flugpfad .....	110

Abbildung 4-46 - Erkennungsdauern in den Aufgaben Warnung und Nachricht .....	111
Abbildung 4-47 - Bewertung der zeitlichen Aspekte .....	111
Abbildung 4-48 - Bewertung der Interventionen in Bezug auf den Arbeitsfluss .....	112
Abbildung 4-49 - Bewertung in Bezug auf Vertrauen und Nützlichkeit.....	112
Abbildung 4-50 - Bewertung im Vergleich zum Versuch ohne Interventionen .....	112
Abbildung 4-51 - Zuweisung von Attributen an die Interventionsgenerierung.....	113
Abbildung 4-52 - Zusammenhang zwischen Interventionsanzahl und mittlerer Delegationszeit.....	114
Abbildung 4-53 - Vergleich der mittleren Delegationszeit zwischen den Versuchen ohne und mit adaptiver Intervention .....	114
Abbildung 4-54 - Erkennungszeiten in der Delegationsaufgabe.....	115
Abbildung 4-55 - Bewertung in Bezug auf zeitliche Aspekte .....	116
Abbildung 4-56 - Bewertung in Bezug auf den Arbeitsfluss .....	116
Abbildung 4-57 - Bewertung in Bezug auf Vertrauen und Nützlichkeit.....	116
Abbildung 4-58 - Bewertung im Vergleich zum Versuch ohne Interventionen .....	117
Abbildung 4-59 - Zuweisung von Attributen an die Interventionsgenerierung.....	117
Abbildung 5-1 - Generische Struktur einer GDTA (Endsley & Jones, 2012).....	122
Abbildung 5-2 - Übergeordnete Struktur der Ziele und Sub-Ziele der GDTA .....	123
Abbildung 5-3 - Sub-Ziel "Identify Threats" .....	124
Abbildung 5-4 - SB-Modell als Wissensbasis für die Interventionen. Annotiert an den Knoten sind die Normierungskonstanten dargestellt (lila).....	125
Abbildung 5-5 - Markierung des aktuellen Flugpfads bei Intervention .....	126
Abbildung 5-6 – Übersicht über die Mission 1.....	130
Abbildung 5-7 – Übersicht über die Mission 2.....	130
Abbildung 5-8 – Übersicht über die Mission 3.....	131
Abbildung 5-9 - Bewertung des Realismus der Missionen (1: „Trifft überhaupt nicht zu“, 5:„Trifft voll und ganz zu“) .....	133
Abbildung 5-10 - Bewertung der Anwendungsfälle Flughöhe, Fluggeschwindigkeit, Flugpfadabweichung und Zeitabweichung .....	133
Abbildung 5-11 - Bewertung der Anwendungsfälle Nachricht, Autopilot und FMS.....	134
Abbildung 5-12 - Bewertung der Anwendungsfälle Neues Bodenziel und Bodengestützte Bedrohung. ....	135
Abbildung 5-13 - Bewertung der Anwendungsfälle Luftgestützte Bedrohung, Niedrige Waffenzahl und Waffe gefeuert.....	136
Abbildung 5-14 - Bewertung des Gesamtsystems über die drei Missionen .....	136
Abbildung 5-15 - Intervention (oben rechts) mit Hinweis auf bodengestützte Bedrohung, während der Proband auf einem kleinen Ausschnitt der Karte die Klassifikation einer SAM-Site durchführt.....	137
Abbildung 5-16 - Reaktion des Probanden auf Intervention durch Anpassung des Kartenausschnitts und Planung der Bekämpfungsaufgabe .....	138
Abbildung 5-17 - Darstellung des MHDD zum Zeitpunkt 12:08:15 bei Proband 2 mit vier aktiven Benachrichtigungen .....	139
Abbildung 5-18 - Verlauf von Flughöhe (blau) und Abweichung im Knoten "AircraftAltitude" (orange). In rot ist der Zeitpunkt der Intervention dargestellt. ....	139

## D. TABELLENVERZEICHNIS

Tabelle 2-1 - Bewertung der Messmethoden von SB .....	29
Tabelle 3-1 - Abweichungsfunktionen der verschiedenen Grunddatentypen .....	58
Tabelle 4-1 - Überblick über die durchgeführten Validierungsstudien mit Teilnehmergruppen, Probanden- und Versuchszahl sowie den abhängigen Variablen.....	60
Tabelle 4-2 - Übersicht über die Textnachrichten in Mission 1. Vorgabe von Fluggeschwindigkeit (VEL/ACC/DECELERATE) und Flughöhe (ALT).....	67
Tabelle 4-3 - Aufgenommene Messdaten im Versuch.....	70
Tabelle 4-4 - Klassifikationsregeln für Reaktionen auf eine Aufgabe .....	71
Tabelle 4-5 - Versuchsteilnehmer. Berechtigungen: 0: Glider; 1: TMG; 2: SEP/MEP Single Pilot; 3: MEP MultiCrew .....	73
Tabelle 4-6 - Vergleich des Durchschnitts, Medians und der Standardabweichung (Std) der Fixationsdauern .....	75
Tabelle 4-7 - Teilnehmergruppe der Kampfflugzeugpiloten mit dem Kürzel „P“ .....	85
Tabelle 4-8 - Mittlere Klassifikationsrate angegeben in Durchschnitt (Standardabweichung). Durchschnittliche Anzahl der Datenpunkte $n = 15612$ . .....	88
Tabelle 4-9 - Parametrisierung der Interventionsgenerierung für Versuch 1 .....	106
Tabelle 4-10 - Parametrisierung der Interventionsgenerierung für Versuch 2 .....	106
Tabelle 4-11 - Durchschnitt und Standardabweichung der Erkennungsdauern für die Verfolgungsaufgaben.....	110
Tabelle 4-12 - Warnungen ( $N_{Warnung} = 20, N_{Nachricht} = 54$ ).....	111
Tabelle 4-13 - Durchschnitt und Standardabweichung der Erkennungszeiten für die Delegationsaufgabe ( $N_{keine Int}=275, N_{Int}=290$ ).....	115
Tabelle 5-1 - Interventionen für das Ziel „Maintain desired flight path“ .....	126
Tabelle 5-2 - Interventionen für das Ziel „Manage Automation“ .....	127
Tabelle 5-3 - Interventionen für das Ziel "Assess ground threats" .....	128
Tabelle 5-4 - Interventionen für das Ziel "Assess air threats" .....	128
Tabelle 5-5 - Interventionen für das Ziel " Optimally assign mission tasks " .....	129
Tabelle 5-6 - Alter und Flugstunden der Teilnehmer .....	132

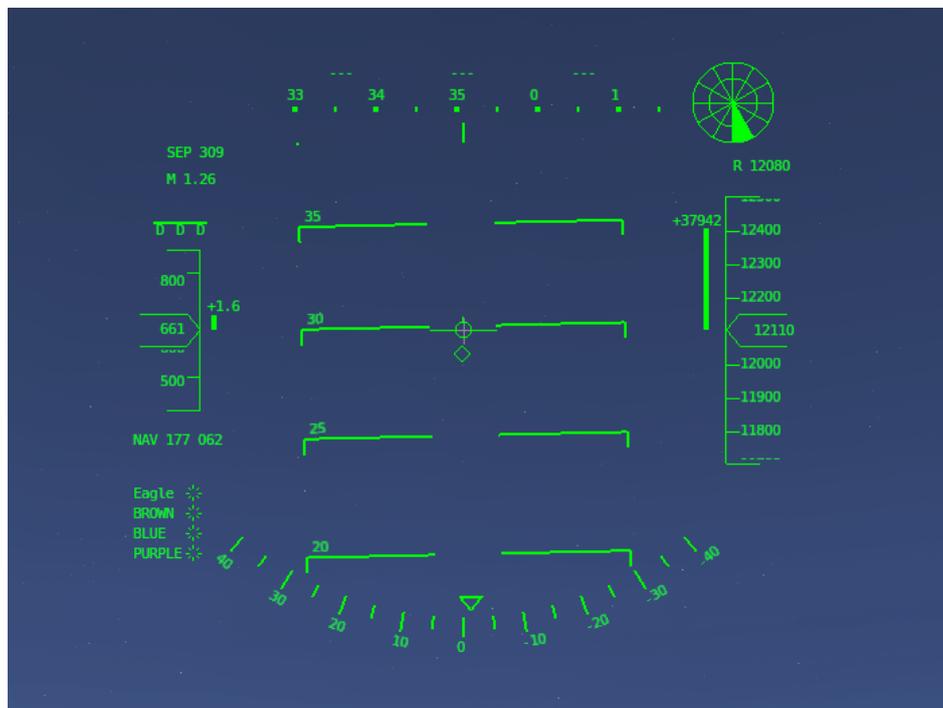
## E. ANHANG

### E.1 BEDIEN- UND ANZEIGEELEMENTE

In diesem Abschnitt werden alle relevanten Anzeige- und Bedienoberflächen des Cockpitsimulators beschrieben. Der Simulator wurde im Rahmen von Forschungsprojekten entwickelt und basierend neben dieser Arbeit auch maßgeblich auf den Entwicklungen von Lindner (2022), Heilemann (2022) und Mund (2023).

#### Head-Up-Display

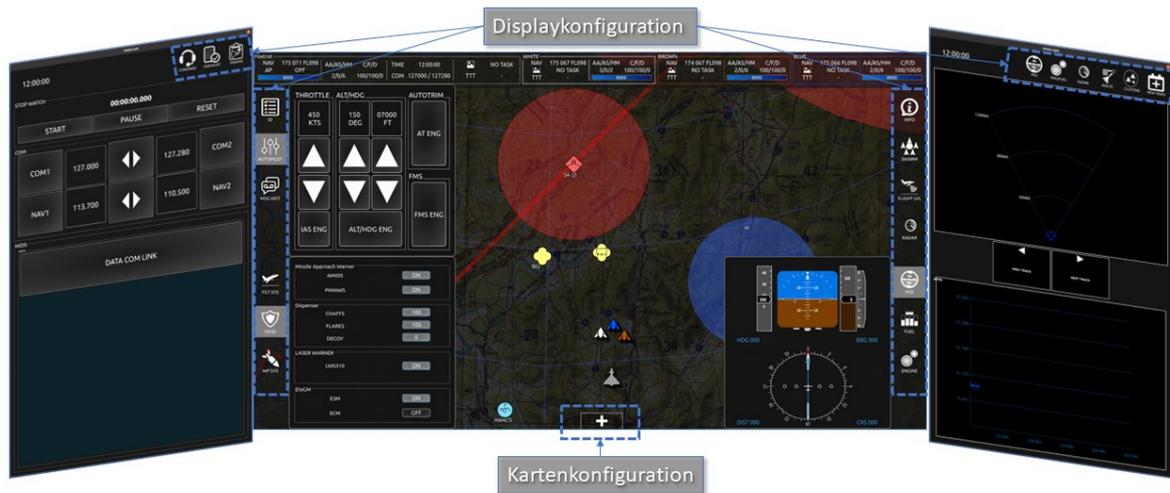
Das Head-Up-Display ist in Appendix Abbildung 1 dargestellt und beinhaltet die wichtigsten Systeminformationen des Flugzeugs. Dazu gehören Flughöhe, Steigrate, Fluggeschwindigkeit, Machzahl, Kurs, Rollwinkel, Nickwinkel, Autopilotenkonfiguration, Radarhöhe, Radarwarner sowie Missionsdaten wie einen Early/Late-Indikator oder eine Annotation der Position anderer Teammitglieder. Im HUD sind außerdem Assistenzfunktionen zur Durchführung von Bekämpfung implementiert.



Appendix Abbildung 1 – Head-Up Display

#### Head-Down-Displays

Das Cockpitlayout besteht aus einem zentralen und zwei seitlichen Touchscreens. Im Zentrum liegt das *Multi-functional Head-Down Displays (MHDD)* mit der taktischen Karte, welche das zentrale Element des Piloten zur Navigation, Auftragsdelegation und Orientierung in der Mission ist. Appendix Abbildung 2 zeigt das Display, an dessen Rändern links und rechts Knöpfe zum Ein- und Ausblenden von Subsystemoberflächen liegen. Am oberen Rand gibt es eine Statusanzeige, die die wichtigsten Informationen über das eigene Flugzeug sowie die unbemannten Plattformen zusammenfasst.

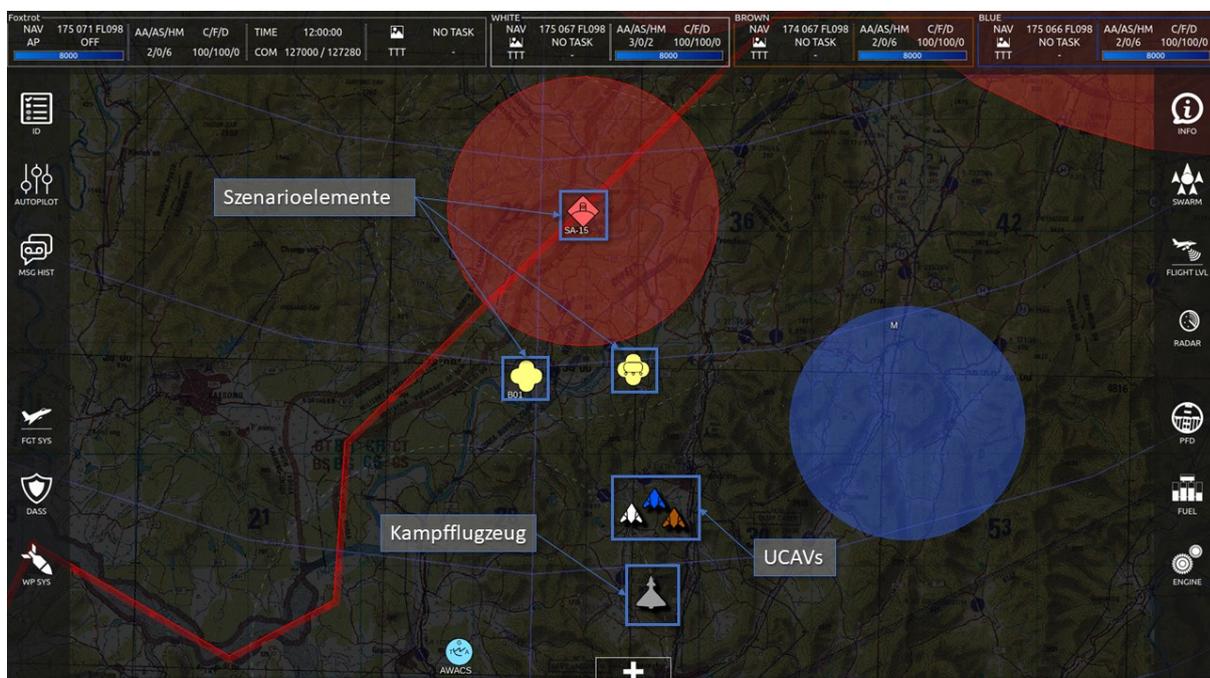


Appendix Abbildung 2 - MHDD mit eingblendeten Autopiloten, DASS und PFD

Rechts und links vom MHDD liegen die *Side Head-Down Displays* (SHDD), deren Anzeige konfigurierbar ist und beispielsweise Kommunikations- oder Treibstoffinformationen bereitstellt.

### Taktische Karte

Die taktische Karte in Appendix Abbildung 3 visualisiert die Positionen der eigenen Luftfahrzeuge sowie bekannte taktische Objekte, wie Gebäude, Fahrzeuge oder Flugabwehrstellungen. Sie lässt sich konfigurieren und kann mit unterschiedlichen Kartendaten, Missionsdaten oder Navigationsmodi betrieben werden.



Appendix Abbildung 3 - Taktische Karte mit markiertem Kampfflugzeug, UCAVs und Szenarioelementen.

Die Appendix Tabelle 1 zeigt die unterschiedlichen Symbole von Gebäuden, Flugzeugen, Fahrzeugen oder Flugabwehrstellungen je nach Klassifikation in unbekannt, freundlich, neutral oder feindlich.

Element	Unbekannt	Friendly	Neutral	Foe
Gebäude				
Flugzeug				
Fahrzeug				
Flugabwehr- stellung	-		-	

Appendix Tabelle 1 - Darstellung von Gebäude, Flugzeug, Fahrzeug und Flugabwehrstellungen in der taktischen Karte

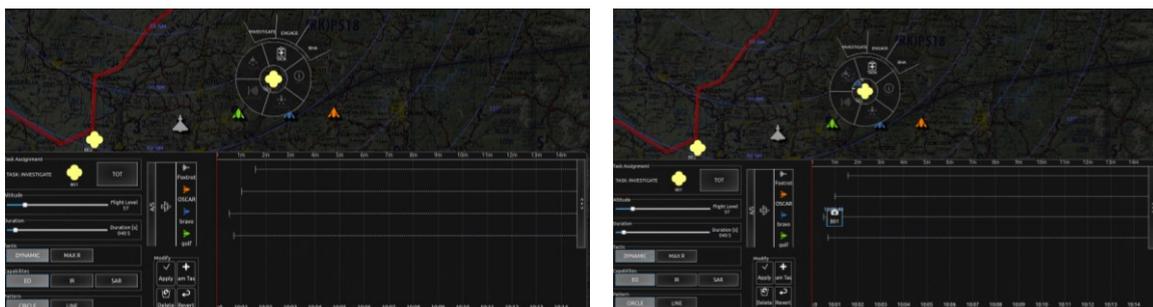
## Missionsplanung

Die Missionsplanung besteht im Wesentlichen aus der Zuweisung von Aufträgen an die unbemannten sowie die eigenen Plattformen in einer Zeitleiste der Mission. Die zentrale Interaktion startet hier in der taktischen Karte durch Anwählen eines Objekts (siehe Appendix Abbildung 4, links) und der Auswahl eines Auftrags (siehe Appendix Abbildung 4, rechts).



Appendix Abbildung 4 - Kontextmenü nach Auswählen eines taktischen Objekts (links) und Kontextmenü nach Auswählen des „Task“-Buttons (rechts).

Der ausgewählte Auftrag wird dann in einer Zeitleiste einem UCAV oder der eigenen Plattform zugewiesen. Die grauen Linien visualisieren den möglichen Zeitraum, in dem ein Auftrag ausgeführt werden kann.

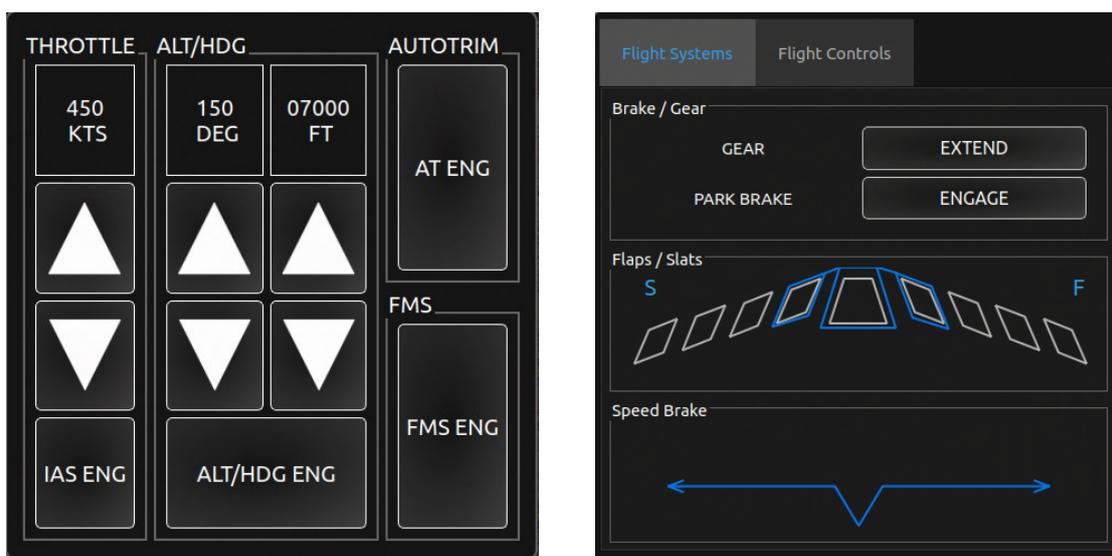


Appendix Abbildung 5 – Interaktion zur Zuweisung eines Auftrags an das blaue UCAV.

Mit diesem Dialog lässt sich zeitleistenbasiert ein Missionsplan anlegen, der die Aufträge der UCAVs im Rahmen der Mission beschreibt. In den Versuchen wurden unterschiedliche Entwicklungsstände des Simulators verwendet, weshalb beispielsweise in den Versuchen im Abschnitt 4.3 eine vertikale Missionsplanungsoberfläche auf dem seitlichen Display verwendet wurde. Die Funktionalität hat sich aber nicht grundlegend verändert, weshalb an dieser Stelle auf eine detaillierte Beschreibung verzichtet wird.

## Autopilot

Der Autopilot (AP) wird über die links in Appendix Abbildung 6 dargestellte Oberfläche bedient. Unter Vernachlässigung der Autotrimmung, welche keinen dauerhaften Modus darstellt, bietet der AP es folgende Modi: Autothrottle, ALT/HDG sowie FMS.



Appendix Abbildung 6 - Bediendisplay für den Autopiloten (links) sowie Flugzeugsysteme (rechts)

Bei aktivierten Flight Management System (FMS) flog das Flugzeug vorher geplante Routenpunkte ab, womit eine Teilautomatisierung von Missionaufgaben möglich war. Autothrottle ermöglichte das Einstellen und Halten einer vorgegebenen Fluggeschwindigkeit, wobei die Maximalgeschwindigkeit 500 Knoten beträgt. ALT/HDG ist ein kombinierter Modus zum Einnehmen und Halten von Flughöhe und Kurs, welcher nur kombiniert aktiviert werden kann.

## Systemmanagement

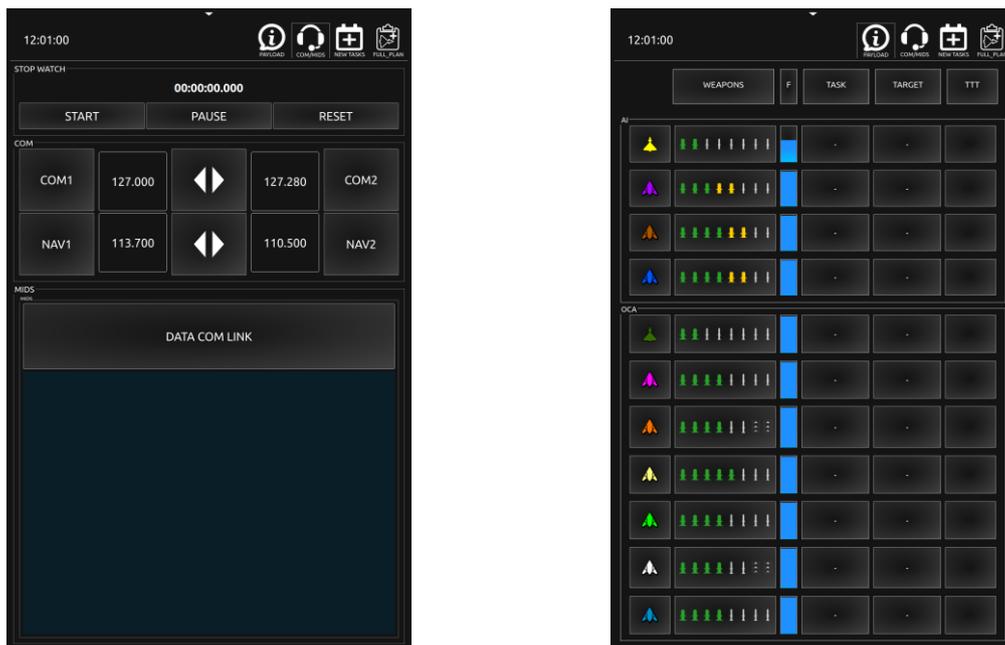
Das Systemmanagement ist in der Cockpitsimulation vereinfacht umgesetzt und besteht im Wesentlichen aus Fahrwerk- und Parkbremseneinstellungen, wie in der Bedienoberfläche dargestellt in Appendix Abbildung 6 ersichtlich ist. Neben dieser Bedienoberfläche gibt es eine Anzeige von

Treibstoff und Hydraulikzustand des Flugzeugs. In den Versuchen war diese Information jedoch nicht relevant.

Die Waffen, der verbleibende Treibstoff der Teammitglieder, die aktuelle Aufgabe, das Ziel sowie die verbleibende Auftragszeit werden in einer Übersichtsanzeige auf einem seitlichen Display angezeigt (siehe Appendix Abbildung 7 rechts).

## Kommunikation

Die Bedienung der Kommunikationssysteme geschieht über die in Abbildung Appendix Abbildung 7 links dargestellte Benutzeroberfläche. Dort können einerseits Funkfrequenzen eingestellt und andererseits Textnachrichten gelesen werden.



Appendix Abbildung 7 - Bedienoberfläche des Kommunikationssystems (links) und Team Ressourcenanzeige (rechts)

## Sensormanagement

Für die Sensorklassifikationsaufgaben in den Versuchen wurden von den UCAVs Bilder von Zielen gemacht, welche im Nachgang in valide und nicht valide Ziele klassifiziert werden mussten („Hot“ & „Cold“). Die Appendix Abbildung 8 zeigt die Oberfläche des Cockpits zur Durchführung dieser Aufgabe. Links im Bild ist eine Liste mit dem Namen des taktischen Objekts und der Anzahl der verfügbaren Bilder, welche nach Anwählen der Zeile in der rechten Spalte untereinander als Thumbnails visualisiert werden. Zusätzlich zum ausgewählten Sensorbild werden im Zentrum noch Metadaten angezeigt. Eine Klassifikation kann dann mit den beiden Knöpfen am rechten Rand des Sensorbilds durchgeführt werden und verändert den Klassifikationsstatus, der in der taktischen Karte angezeigt wird. Eine Klassifikation zu „HOT“ ändert beispielsweise den Status eines Gebäudes auf feindlich.



Appendix Abbildung 8 - Oberfläche zur Durchführung einer Zielklassifikation

## E.2 VALIDIERUNG DER AUFMERKSAMKEITSMESSUNG

### E.2.1. MISSION 1

Zeitpunkt	Skriptangabe	Zeitpunkt	Skriptangabe
14:00:01	MIDS FALCON CLEARED FOR T/O RWY 08 R	14:19:20	MIDS DECELERATE 340 kts ALT 4500 ft
14:01:45	MIDS FALCON CLB 4500 ft VEL 350 kts	14:19:35	WIND ALTO03870
14:02:07	WIND VELO301	14:19:48	WIND VELO388
14:02:29	WIND VELO399	14:20:01	WIND VELO388
14:02:51	WIND VELO301	14:20:14	WIND ALTO05130
14:03:13	WIND ALTO05130	14:20:27	WIND ALTO05130
14:03:35	WIND ALTO05130	14:20:40	WIND ALTO05130
14:03:57	WIND ALTO05130	14:20:53	WIND VELO293
14:04:19	WIND VELO399	14:21:06	WIND ALTO05130
14:04:41	WIND ALTO05130	14:21:19	WIND VELO293
14:05:03	WIND VELO301	14:21:32	WIND VELO388
14:05:25	WIND VELO399	14:21:45	WIND ALTO05130
14:05:47	WIND VELO399	14:21:58	WIND VELO317
14:06:09	WIND VELO399	14:22:11	WIND VELO317
14:06:31	WIND ALTO05130	14:22:24	WIND VELO317
14:06:53	WIND ALTO05130	14:22:37	WIND VELO364
14:07:15	WIND VELO399	14:22:50	WIND VELO317
14:07:25	MIDS ACCELERATE 390kts ALT 4500 ft	14:23:03	WIND ALTO04815
14:07:37	WIND ALTO03870	14:23:16	WIND VELO317
14:07:59	WIND ALTO05130	14:23:29	WIND VELO364
14:08:21	WIND ALTO05130	14:23:42	WIND ALTO04815
14:08:43	WIND VELO336	14:23:52	MIDS ACCELERATE 360kts ALT 4500 ft
14:09:05	WIND VELO445	14:23:55	WIND ALTO04815
14:09:27	WIND VELO336	14:24:08	WIND VELO386
14:09:49	WIND ALTO04185	14:24:21	WIND ALTO04815
14:10:11	WIND ALTO04815	14:24:34	WIND VELO335
14:10:33	WIND ALTO04185	14:24:47	WIND ALTO04815
14:10:55	WIND ALTO04815	14:25:00	WIND ALTO04185
14:11:17	WIND VELO363	14:25:13	WIND VELO386
14:11:39	WIND VELO363	14:25:26	WIND VELO335
14:12:01	WIND VELO363	14:25:39	WIND ALTO04815
14:12:23	WIND ALTO04815	14:25:52	WIND ALTO04815
14:12:45	WIND ALTO04815	14:26:05	WIND VELO335
14:12:55	MIDS DECELERATE 360 kts ALT 4500 ft	14:26:22	WIND ALTO04815
14:13:15	WIND ALTO04185	14:26:39	WIND VELO335
14:13:37	WIND VELO386	14:26:56	WIND VELO386
14:13:59	WIND ALTO04815	14:27:13	WIND VELO335
14:14:21	WIND VELO335	14:27:30	WIND VELO386
14:14:43	WIND VELO386	14:27:47	WIND ALTO04185
14:15:05	WIND ALTO04815	14:28:04	WIND ALTO04185
14:15:27	WIND ALTO04815	14:28:21	WIND VELO386
14:15:49	WIND ALTO04500	14:28:38	WIND ALTO04815
14:16:11	WIND ALTO04500	14:28:55	WIND ALTO04185
14:16:33	WIND VELO386	14:29:12	WIND ALTO04185
14:17:13	WIND ALTO05130	14:29:29	WIND VELO335
14:17:26	WIND VELO411	14:29:46	WIND VELO335
14:17:39	WIND VELO411	14:30:03	WIND ALTO04815
14:17:52	WIND ALTO03870	14:30:20	WIND ALTO04185
14:18:05	WIND VELO310	14:30:37	WIND VELO386
14:18:18	WIND VELO310	14:30:54	WIND VELO386
14:18:31	WIND VELO310		
14:18:44	WIND ALTO05130		
14:18:57	WIND ALTO05130		
14:19:10	WIND ALTO03870		

Appendix Tabelle 2 - Detaillierter Ablauf der Mission 1. „WIND VEL/ALT<X>“ löst eine Böe aus, die das Flugzeug auf die Höhe/Geschwindigkeit X setzt. „MIDS <TEXT>“ erzeugt eine Nachricht mit dem Inhalt TEXT.

## E.2.2. MISSION 2

Zeitpunkt	Skriptangabe
14:00:10.000	MIDS FALCON CLEARED FOR T/O RWY 36
14:01:55.000	MIDS FALCON ACTIVATE AUTOPILOT CLIMB 8000 ft ACCELERATE 500 kts TURN RIGHT HDG 120
14:02:22.000	Set Known SASE1 SAW1
14:02:49.000	Set Known SASW1 BUNW1
14:03:16.000	Set Known SASE2 BUW1
14:03:43.000	MIDS FALCON TURN RIGHT HDG 300 CLIMB 9000 ft
14:04:10.000	Set Known OILDEPOT1 AMBULANCE VES1 BUS1
14:04:37.000	Set Known SAW3 SAW2
14:05:04.000	Set Known TENT1 SASW2
14:05:31.000	MIDS FALCON TURN RIGHT HDG 320 DESCENT 7500
14:05:58.000	Set Known BUNW2
14:06:25.000	Set Known SASE3
14:06:52.000	Set Known VEW1
14:07:19.000	Set Known SANW1
14:07:46.000	MIDS FALCON TURN RIGHT HDG 140 CLIMB 9000 ft
14:08:13.000	Set Known CHURCH1
14:08:40.000	Set Dead AMBULANCE
14:09:02.000	Set Known POWERSTATION1
14:09:21.000	Set Known SASW3
14:09:40.000	MIDS FALCON TURN LEFT HDG 110 CLIMB 10000
14:09:59.000	Set Known SAS1
14:10:18.000	MIDS FALCON DECELERATE 450 kts TURN RIGHT HDG 290
14:10:37.000	Set Known VENW1
14:10:56.000	MIDS FALCON TURN RIGHT HDG 330
14:11:15.000	Set Known BUSW1 SASW4
14:11:34.000	Set Known SAS2
14:11:53.000	Set Known SANW2
14:12:12.000	MIDS FALCON TURN RIGHT HDG 140 DESCENT 8000 ft
14:12:31.000	Set Known BUS2
14:12:50.000	Set Known OILDEPOT2 AMBULANCE2 Set VENW2
14:13:09.000	Set Known SAW4
14:13:28.000	MIDS FALCON TURN LEFT HEADING 120 ACCELERATE 500 kts
14:13:47.000	Set Known BUSW2
14:14:06.000	Set Known TRUCK2 TRUCK1 SANW3
14:14:25.000	Set Known BUSE1 SASW5
14:14:44.000	MIDS FALCON TURN LEFT HDG 310 DECELERATE 400 kts
14:15:03.000	Set Known SASW6
14:15:22.000	Set Known SAW5
14:15:41.000	Set Known BUW2 POWERSTATION2
14:16:00.000	MIDS FALCON TURN RIGHT HDG 330 VEL 400 kts ALT 8000 ft
14:16:19.000	Set Known SASE4
14:16:38.000	Set Known VESW1
14:16:57.000	FAIL AP_IAS
14:17:16.000	MIDS FALCON TURN RIGHT HDG 150 VEL 350 kts 8000 ft
14:17:35.000	Set Known SASE5 FACTORY1 TANK
14:17:54.000	Set Known BUSW3 OILDEPOT3
14:18:13.000	Set Known SAW6 VESE1
14:18:32.000	MIDS FALCON TURN LEFT HEADING 130 DESCENT 5500 ft VEL 350 kts
14:18:52.000	FIGHTER WIND VEL0390
14:19:12.000	FIGHTER WIND VEL0380
14:19:32.000	FIGHTER WIND VEL0390
14:19:52.000	FIGHTER WIND ALT004950
14:20:12.000	MIDS FALCON RETURN TO HOMEBASE. DESCENT 5000 ft VEL 350 kts
14:20:32.000	FIGHTER WIND ALT004000
14:20:52.000	FIGHTER WIND 0315
14:21:12.000	FIGHTER WIND VEL0380
14:21:32.000	FIGHTER WIND ALT004900
14:21:52.000	FIGHTER WIND ALT004900
14:22:12.000	FIGHTER WIND ALT004100
14:22:32.000	FIGHTER WIND VEL0380
14:22:52.000	FIGHTER WIND VEL0390

Appendix Tabelle 3 - Detaillierter Ablauf der Mission 2. Ausgewerteter Zeitraum ist blau hinterlegt.

## E.2.3. ERGEBNISSE

Radius	RP	FP	FN	Prazision	Sensitivitat
.1°	241	34	607	0.87636364	0.28419811
0.5°	323	61	498	0.84114583	0.39342266
1.0°	438	88	356	0.83269962	0.55163728
1.5°	524	114	244	0.82131661	0.68229167
2.0°	571	136	175	0.80763791	0.76541555
2.5°	615	147	120	0.80708661	0.83673469
3.0°	618	152	112	0.8025974	0.84657534
4.0°	616	156	110	0.79792746	0.84848485
5.0°	614	158	110	0.79533679	0.8480663
6.0°	612	162	108	0.79069767	0.85
7.0°	618	164	100	0.79028133	0.86072423

Appendix Tabelle 4 - Richtig-positiv, falsch-positiv und richtig-negative Ergebnisse

## E.3 VALIDIERUNG DER SB-MESSUNG

### E.3.1. SART

**Situation Awareness Versuchsreihe**

ID: \_\_\_\_\_

Versuch: \_\_\_\_\_

**1. Wie schnell konnte sich die Situation verandern?**  
 War sie instabil und konnte sich plotzlich verandern (7) oder war sie stabil und voraussehbar (1)?

1  2  3  4  5  6  7

**2. Wie kompliziert war die Situation?**  
 War sie komplex mit vielen zusammenhangende Komponenten (7) oder war sie einfach und durchschaubar (1)?

1  2  3  4  5  6  7

**3. Wie viele Faktoren anderten sich in der Situation?**  
 anderten sich sehr viele (7) oder kaum welche (1)?

1  2  3  4  5  6  7

**4. Wie Aufgeregt waren Sie in der Situation?**  
 Waren Sie alarmiert und bereit zum Handeln (7) oder hatten Sie eine geringe Alarmbereitschaft (1)?

1  2  3  4  5  6  7

**5. Wie stark haben Sie sich auf die Situation konzentriert?**  
 Haben Sie sich auf viele Aspekte der Situation konzentriert (7) oder waren sie nur auf einen fokussiert (1)?

1  2  3  4  5  6  7

**Situation Awareness Versuchsreihe**

ID: \_\_\_\_\_

Versuch: \_\_\_\_\_

**6. Wie hoch war Ihre Aufmerksamkeit in der Situation?**

Halten Sie viele Aspekte der Situation wahrgenommen (7) oder nur einen (1)?

- 1     2     3     4     5     6     7

**7. Wie war Ihre mentale Auslastung?**

Halten Sie genug Kapazität, um sich mit vielen Faktoren zu befassen (7), oder waren Sie vollständig ausgelastet (1)?

- 1     2     3     4     5     6     7

**8. Wie viele Informationen haben Sie über die Situation bekommen?**

Halten Sie sehr viele Kenntnisse bekommen und verstanden (7) oder sehr wenige (1)?

- 1     2     3     4     5     6     7

**9. Wie vertraut waren Sie mit der Situation?**

Halten Sie schon viele relevante Erfahrungen (7) oder war für Sie die Situation neu (1)?

- 1     2     3     4     5     6     7

Appendix Abbildung 9 - SART Fragebogen

## E.4 VALIDIERUNG INTERVENTIONSGENERIERUNG

### E.4.1. MISSION 1

Zeitpunkt	Skriptangabe	Zeitpunkt	Skriptangabe
14:06:53.000	WIND ALT005710	14:21:32.000	WIND VEL0388
14:07:15.000	WIND VEL0360	14:21:45.000	WIND ALT005130
14:07:25.000	MIDS ACCELERATE 450kts 4500 ft	14:21:58.000	WIND VEL0317
14:07:37.000	WIND ALT004000	14:22:11.000	WIND VEL0317
14:07:59.000	WIND ALT004910	14:22:24.000	WIND VEL0317
14:08:21.000	WIND ALT004910	14:22:37.000	WIND VEL0364
14:08:43.000	WIND VEL0420	14:22:50.000	WIND VEL0317
14:09:05.000	WIND VEL0410	14:23:03.000	WIND ALT004815
14:09:27.000	WIND VEL0400	14:23:16.000	WIND VEL0317
14:09:49.000	WIND ALT004910	14:23:29.000	WIND VEL0364
14:10:11.000	WIND ALT004910	14:23:42.000	WIND ALT004815
14:10:33.000	WIND ALT004910	14:23:52.000	MIDS ACCELERATE 400kts 5000 ft
14:10:55.000	WIND ALT004815	14:23:55.000	WIND ALT004815
14:11:17.000	WIND VEL0410	14:24:08.000	WIND VEL0386
14:11:39.000	WIND VEL0415	14:24:21.000	WIND ALT004815
14:12:01.000	WIND VEL0415	14:24:34.000	WIND VEL0335
14:12:23.000	WIND ALT004910	14:24:47.000	WIND ALT004815
14:12:45.000	WIND ALT004815	14:25:00.000	WIND ALT004185
14:12:55.000	MIDS DECELERATE 360 kts 5000 ft	14:25:13.000	WIND VEL0386
14:13:15.000	WIND ALT004600	14:25:26.000	WIND VEL0335
14:13:37.000	WIND VEL0386	14:25:39.000	WIND ALT004815
14:13:59.000	WIND ALT004815	14:25:52.000	WIND ALT004815
14:14:21.000	WIND VEL0335	14:26:05.000	WIND VEL0335
14:14:43.000	WIND VEL0386	14:26:22.000	WIND ALT004815
14:15:05.000	WIND ALT004815	14:26:39.000	WIND VEL0335
14:15:27.000	WIND ALT004815	14:26:56.000	WIND VEL0386
14:15:49.000	WIND ALT004500	14:27:13.000	WIND VEL0335
14:16:11.000	WIND ALT004500	14:27:30.000	WIND VEL0386
14:16:33.000	WIND VEL0400	14:27:47.000	WIND ALT004185
14:17:13.000	WIND ALT005130	14:28:04.000	WIND ALT004185
14:17:26.000	WIND VEL0411	14:28:21.000	WIND VEL0386
14:17:39.000	WIND VEL0411	14:28:38.000	WIND ALT004815
14:17:52.000	WIND ALT003870	14:28:55.000	WIND ALT004185
14:18:05.000	WIND VEL0310	14:29:12.000	WIND ALT004185
14:18:18.000	WIND VEL0310	14:29:29.000	WIND VEL0335
14:18:31.000	WIND VEL0310	14:29:46.000	WIND VEL0335
14:18:44.000	WIND ALT005130	14:30:03.000	WIND ALT004815
14:18:57.000	WIND ALT005130	14:30:20.000	WIND ALT004185
14:19:10.000	WIND ALT003870	14:30:37.000	WIND VEL0386
14:19:20.000	MIDS DECELERATE 340 kts 4500 ft	14:30:54.000	WIND VEL0386
14:19:35.000	WIND ALT003870	14:29:12.000	WIND ALT004185
14:19:48.000	WIND VEL0388	14:29:29.000	WIND VEL0335
14:20:01.000	WIND VEL0388	14:29:46.000	WIND VEL0335
14:20:14.000	WIND ALT005130		
14:20:27.000	WIND ALT005130		
14:20:40.000	WIND ALT005130		
14:20:53.000	WIND VEL0293		
14:21:06.000	WIND ALT005130		
14:21:19.000	WIND VEL0293		

Appendix Tabelle 5 - Detaillierter Ablauf der Mission 1. „WIND VEL/ALT<X>“ löst eine Böe aus, die das Flugzeug auf die Höhe/Geschwindigkeit X setzt. „MIDS <TEXT>“ erzeugt eine Nachricht mit dem Inhalt TEXT.

## E.4.2. MISSION 2

Zeitpunkt	Skriptangabe
14:01:50.000	Set Known SASE1 Set Known SB-14
14:02:15.000	Set Known SB-14 Set Known BUNW1
14:02:20.000	Set Known SASE2 Set Known BUW1
14:03:00.000	Set Known OILDEPOT1 Set Known AMBULANCE Set Known VES1 Set Known BUS1
14:03:10.000	Set Known BUNW2 Set Known SB-14
14:03:50.000	Set Known TENT1 Set Known SB-14
14:04:10.000	Set Known BUNW2
14:04:13.000	Set Known Building Set Known SB-14
14:04:20.000	Set Known VEW1
14:04:50.000	Set Known SANW1 Set Known SB-14
14:05:05.000	Set Known CHURCH1 Set Known H01
14:05:15.000	Set Known T07 Set Known AMBULANCE
14:05:45.000	Set Known POWERSTATION1 Set Dead CHURCH1 Set Dead OILDEPOT1
14:06:10.000	Set Known SB-14
14:06:20.000	Set Known SB-14
14:06:22.000	Set Known VENW1
14:06:27.000	Set Known BUSW1 Set Known SB-14
14:06:55.000	Set Known SAS2 Set Known Building
14:07:00.000	Set Known SB-14 Set Known Building
14:07:15.000	Set Known BUS2
14:07:40.000	Set Known OILDEPOT2 Set Known AMBULANCE2 Set Known VENW2
14:08:20.000	Set Known SAW4
14:08:40.000	Set Known BUSW2
14:08:59.000	Set Known TRUCK2 Set Known TRUCK1 Set Known SANW3
14:09:10.000	Set Known BUSE1 Set Known SASW5
14:09:27.000	Set Known SB-14
14:09:45.000	Set Known SB-14
14:10:00.000	Set Known BUW2 Set Known POWERSTATION2
14:10:20.000	Set Known SASE4
14:10:50.000	Set Known VESW1
14:11:00.000	Set Known SB-14 Set Known FACTORY1 Set Known TANK
14:11:01.000	Set Known BUSW3 Set Known OILDEPOT3
14:11:07.000	Set Known SAW6
14:01:50.000	Set Known SASE1 Set Known SB-14
14:02:15.000	Set Known SB-14 Set Known BUNW1
14:02:20.000	Set Known SASE2 Set Known BUW1
14:03:00.000	Set Known OILDEPOT1 Set Known AMBULANCE Set Known VES1 Set Known BUS1
14:03:10.000	Set Known BUNW2 Set Known SB-14
14:03:50.000	Set Known TENT1 Set Known SB-14
14:04:10.000	Set Known BUNW2
14:04:13.000	Set Known Building Set Known SB-14
14:04:20.000	Set Known VEW1
14:04:50.000	Set Known SANW1 Set Known SB-14
14:05:05.000	Set Known CHURCH1 Set Known H01
14:05:15.000	Set Known T07 Set Known AMBULANCE
14:05:45.000	Set Known POWERSTATION1 Set Dead CHURCH1 Set Dead OILDEPOT1
14:06:10.000	Set Known SB-14
14:06:20.000	Set Known SB-14
14:06:22.000	Set Known VENW1
14:06:27.000	Set Known BUSW1 Set Known SB-14
14:06:55.000	Set Known SAS2 Set Known Building
14:07:00.000	Set Known SB-14 Set Known Building
14:07:15.000	Set Known BUS2
14:07:40.000	Set Known OILDEPOT2 Set Known AMBULANCE2 Set Known VENW2
14:08:20.000	Set Known SAW4
14:08:40.000	Set Known BUSW2
14:08:59.000	Set Known TRUCK2 Set Known TRUCK1 Set Known SANW3
14:09:10.000	Set Known BUSE1 Set Known SASW5
14:09:27.000	Set Known SB-14
14:09:45.000	Set Known SB-14
14:10:00.000	Set Known BUW2 Set Known POWERSTATION2
14:10:20.000	Set Known SASE4
14:10:50.000	Set Known VESW1
14:11:00.000	Set Known SB-14 Set Known FACTORY1 Set Known TANK
14:11:01.000	Set Known BUSW3 Set Known OILDEPOT3
14:11:07.000	Set Known SAW6

Appendix Tabelle 6 - Detaillierter Ablauf der Mission 2

### E.4.3. FRAGEBOGEN

PILOT ID \_\_\_\_\_

#### 1. Eigenschaften des Assistenzsystems

Bitte geben Sie an, wie stark folgende Aussagen für die Assistenzmeldungen in dieser Mission zutreffen:

	Trifft voll und ganz zu	Trifft überhaupt nicht zu								
1. Die Meldungen kamen zum richtigen Zeitpunkt										
2. Die Frequenz der Meldungen war angemessen										
3. Der Zeitpunkt der Meldungen war nachvollziehbar										
4. Die Meldungen lenkten mich von wichtigeren Aufgaben ab										
5. Die Meldungen waren aufdringlich										
6. Die Meldungen zogen Aufmerksamkeit auf sich										
7. Die Meldungen störten mich										
8. Die Meldungen integrieren sich in meinen Arbeitsablauf										
9. Ich habe die Meldungen nicht beachtet										
10. Ich habe mich auf die Meldungen verlassen										
11. Die Meldungen waren hilfreich										

Bitte geben Sie an, wie stark folgende Aussagen im Vergleich mit dem Versuch ohne Assistenzsystem zutreffen:

	Trifft voll und ganz zu	Trifft überhaupt nicht zu								
1. Meine Leistung hat sich verbessert										
2. Meine Belastung war größer										
3. Meine Frustration war größer										
4. Ich habe mein Verhalten geändert										

Falls zutreffend, in wie fern haben Sie Ihr Verhalten bei der Durchführung der Aufgabe geändert:

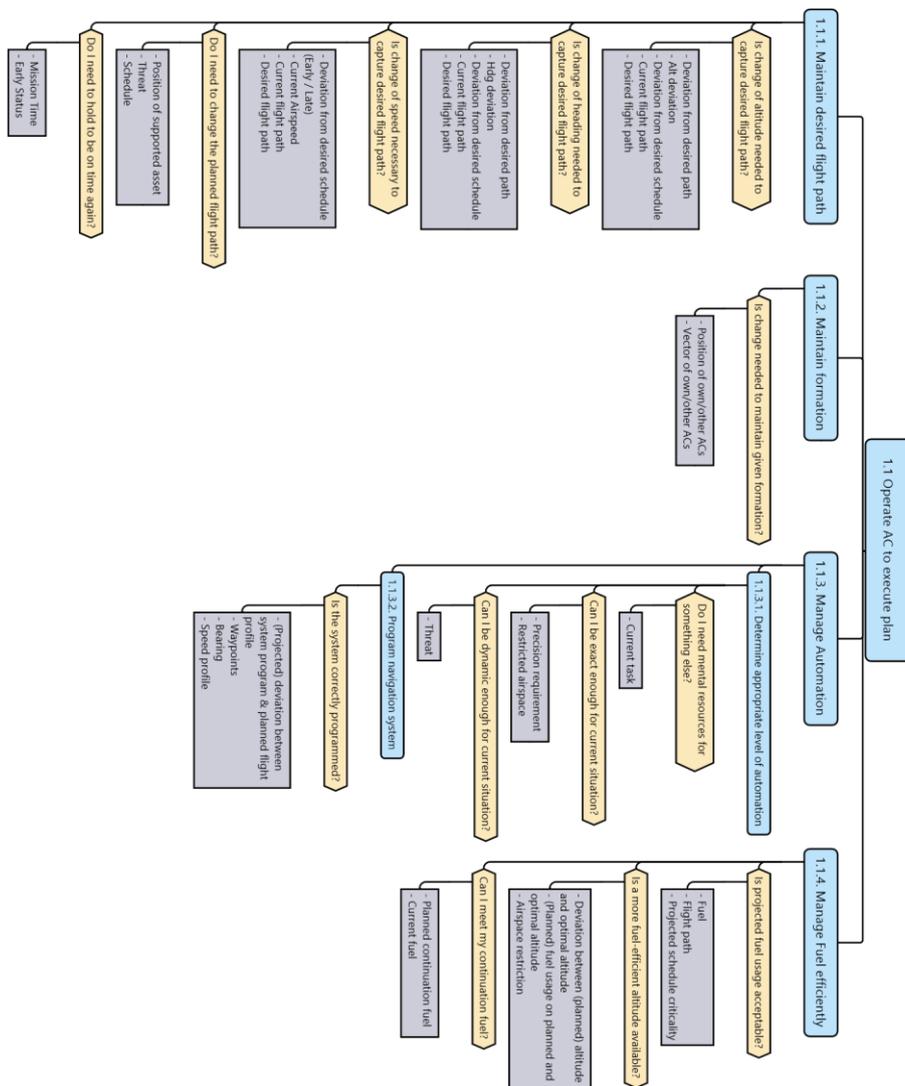
Bitte wählen Sie für jede Kategorie aus, wie stark das jeweilige Attribut auf die Assistentenfunktionalität zutrifft. Die Assistenten verhält sich...

	Verständlich	Nachvollziehbar	Unverständlich	Nicht nachvollziehbar
Verständlich				
Nachvollziehbar				
Vorhersehbar			Unvorhersehbar	
Zuverlässig			Unzuverlässig	
Vorschnell			Zurückhaltend	
Dumm			Schlau	

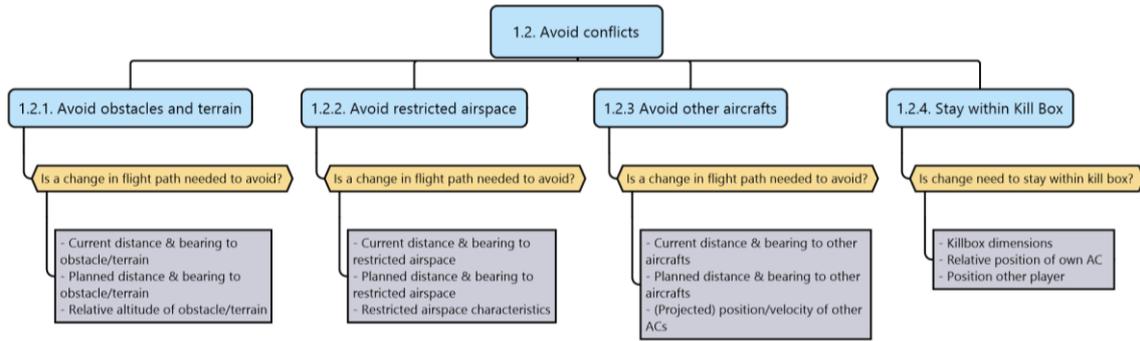
#### 2. Kritik

Was ist verbesserungswürdig?

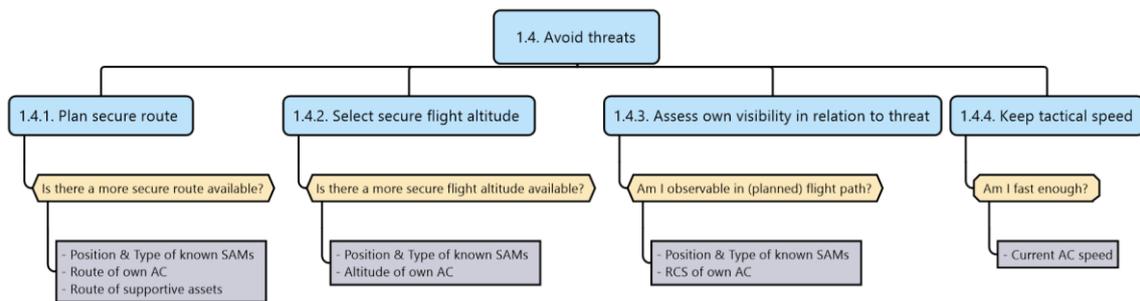
E.5 GDTA



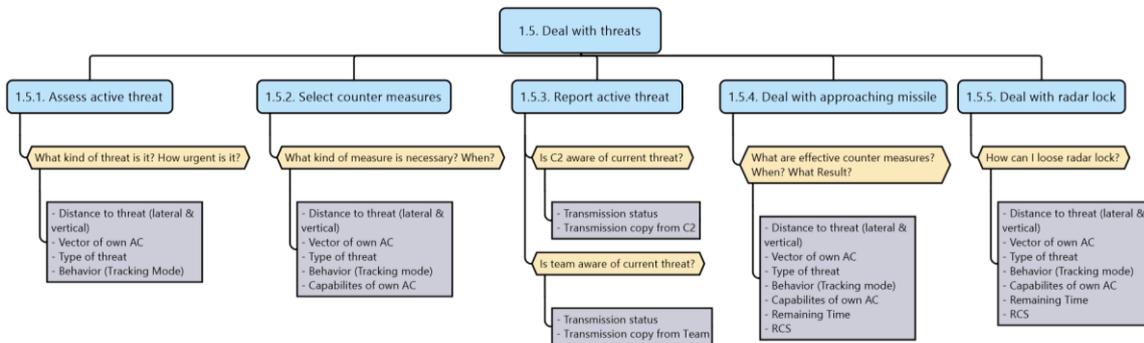
Appendix Abbildung 10 - Subziel "Operate AC to execute plan"



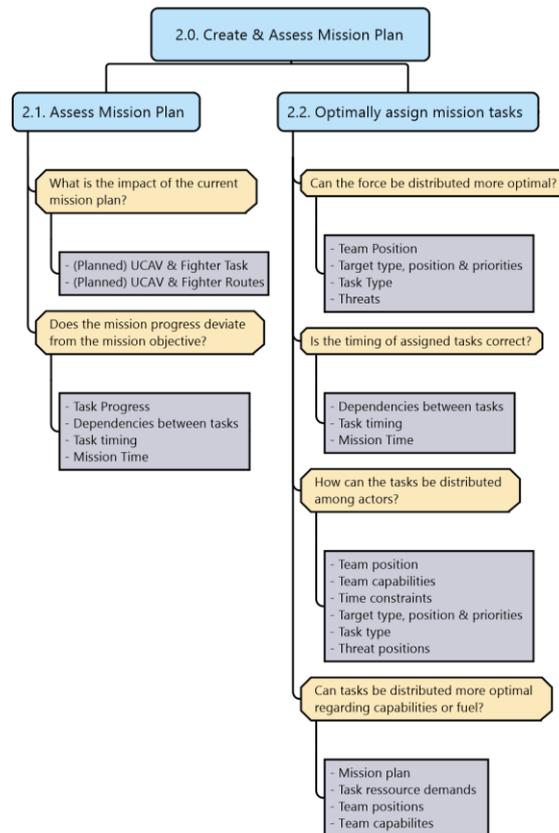
Appendix Abbildung 11 – Subziel „Avoid conflicts“



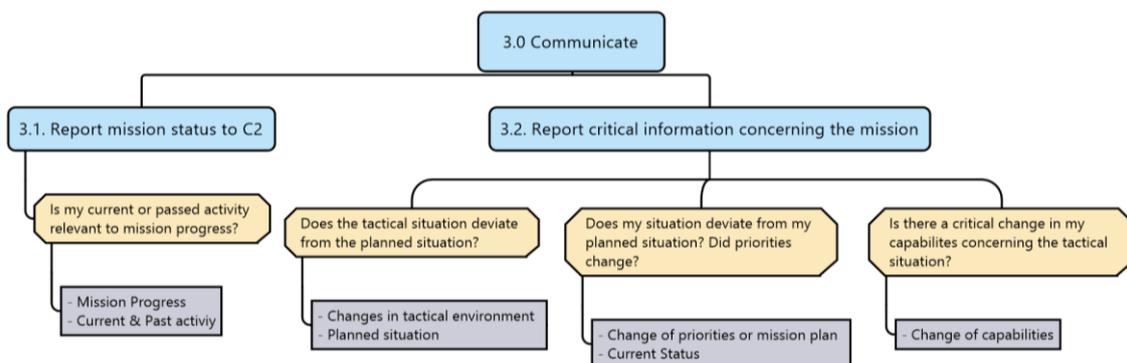
Appendix Abbildung 12 - Subziel "Avoid Threats"



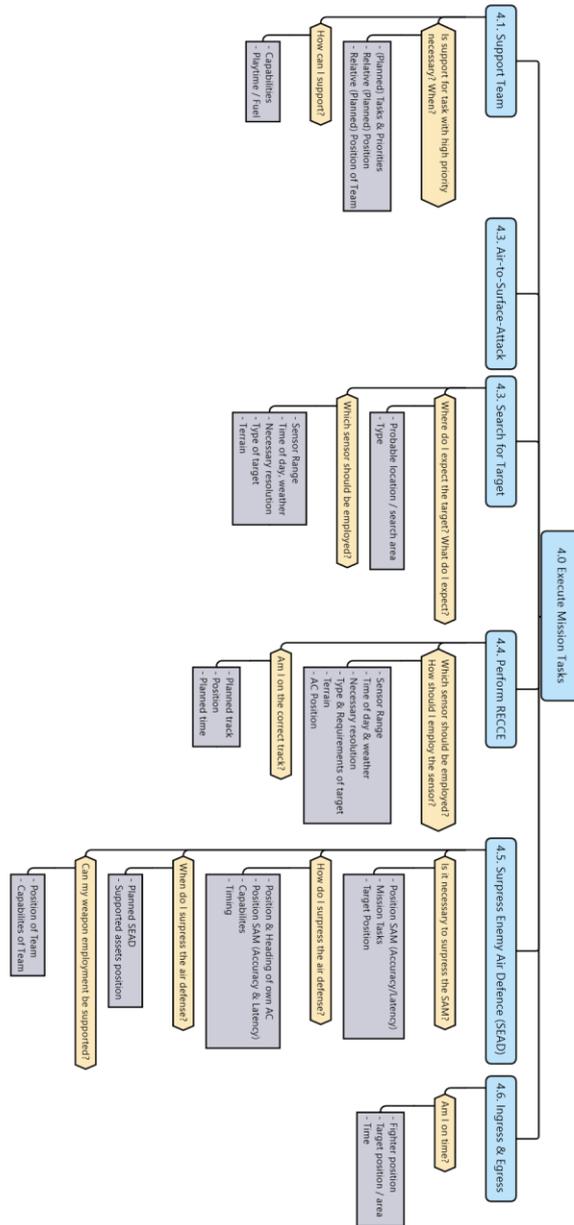
Appendix Abbildung 13 - Subziel "Deal with threats"



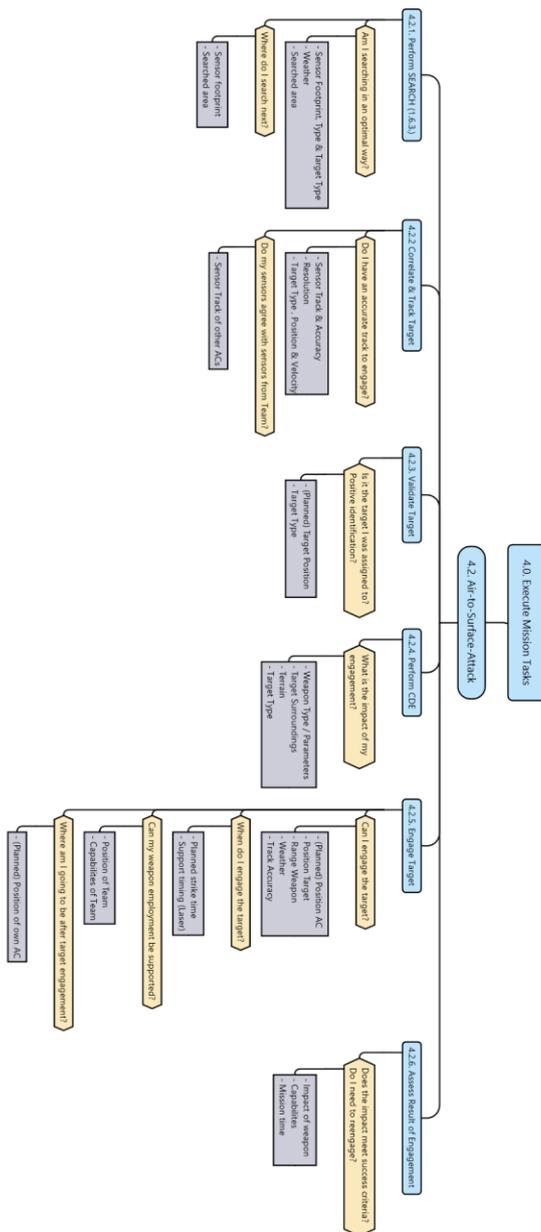
Appendix Abbildung 14 - Subziel "Create & Assess Mission Plan"



Appendix Abbildung 15 - Subziel „Communicate“



Appendix Abbildung 16 - Subziel "Execute Mission Plan with Fighter AC".



Appendix Abbildung 17 - Subziel "Perform A/I Objective"

## E.6 VALIDIERUNG MUM-T GESAMTSYSTEM

### E.6.1. MISSION 1

Zeitpunkt / Wegpunkt	Nachricht
12:00:01.000	MIDS CLEARED FOR TAKE OFF. MANUAL FLIGHT. FL 20.
P0	MIDS MANUAL FLIGHT. FL 40.
P1	MIDS MANUAL FLIGHT. FL 60.
P2	MIDS AUTOPILOT FL 80.
P3	MIDS AUTOPILOT FL 90.
P4	MIDS AUTOPILOT FL 100.
P5	MIDS FMS
P8	MIDS MANUAL FLIGHT. FL 50.
P9	MIDS MANUAL FLIGHT. FL 40.
P10	MIDS CLEARED FOR LANDING

Appendix Tabelle 7 - Anweisungen über Flughöhe und die Verwendung von Automation in der Mission 1

### E.6.2. MISSION 2

Objekt	Zeitpunkt Auftauchen	Penetrationszeit
Air3	12:01:00	12:04:53
Air2	12:02:10	12:05:13
Air5	12:03:00	12:04:40
Air1	12:05:30	12:09:45
Air4	12:07:30	12:12:15
S2_SB-15	12:08:00	-
S1_SB-15	12:09:10	-

Appendix Tabelle 8 - Zeiten kritischer Ereignisse in der Mission 2

## E.6.3. FRAGEBOGEN

### Assistenzfall Altitude

Beantworten Sie die folgenden Fragen für den Assistenzfall "Altitude".

Beispiel: "FLIGHT LEVEL SIX ZERO TARGET ALTITUDE IS AT NINE ZERO"

Die Meldung wurde ausgelöst	<input type="radio"/> Ja <input type="radio"/> Nein	
Die Meldung war <b>hilfreich</b>	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldung kam zum <b>richtigen Zeitpunkt</b>	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die <b>Frequenz</b> der Meldung war <b>angemessen</b>	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldung lenkte meine <b>Aufmerksamkeit</b> auf <b>relevante Informationen</b>	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldung lenkte meine <b>Aufmerksamkeit</b> auf Informationen, die mir <b>nicht bewusst</b> waren...	<input type="radio"/>	<input type="radio"/> Keine Angabe
Ich konnte mich darauf <b>verlassen</b> , dass mich die Meldung auf <b>wichtige Informationen</b> hinweist.	<input type="radio"/>	<input type="radio"/> Keine Angabe

Was war an der Gestaltung der Meldung gut?

Was ist an der Gestaltung der Meldung verbesserungswürdig?

### Allgemeine Bewertung

1. Bitte geben Sie an, wie stark folgende Aussagen für die Assistenzmeldungen in dieser Mission zutreffen:

Meine Leistung hat sich durch die Assistenz insgesamt verbessert	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Assistenz war insgesamt hilfreich, um den Überblick zu behalten	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die auditiven Meldungen waren insgesamt nützlich	<input type="radio"/>	<input type="radio"/> Keine Angabe
Der Inhalt der auditiven Nachrichten war nützlich	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldungen auf der takt. Karte waren insgesamt nützlich.	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldungen lenkten meine Aufmerksamkeit auf Informationen, die mir nicht bewusst waren.	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldungen integrierten sich gut in meinen Arbeitsablauf.	<input type="radio"/>	<input type="radio"/> Keine Angabe
Die Meldungen lenkten von wichtigeren Aufgaben ab.	<input type="radio"/>	<input type="radio"/> Keine Angabe
Ich konnte mich auf die Meldungen verlassen.	<input type="radio"/>	<input type="radio"/> Keine Angabe

2. Bitte wählen Sie für jede Kategorie aus, wie stark das jeweilige Attribut auf die Assistenzfunktionalität zutrifft. Die Assistenz verhält sich...

Unverständlich	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	Verständlich
Nachvollziehbar	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	Nicht nachvollziehbar
Vorhersehbar	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	Unvorhersehbar
Zuverlässig	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	Unzuverlässig
Dumm	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	Schlau

3. Sonstige Anmerkungen

Was ist am Training der Meldungen verbesserungswürdig?

Was ist an der Art der Intervention (z.B. Modalität, Gestaltung) der Interventionen verbesserungswürdig?

Welche anderen Anwendungsfälle wären sinnvoll?

4. Geben Sie bitte an, inwiefern die folgenden Aussagen auf den Versuch zutreffen:

Trifft überhaupt nicht zu Trifft voll und ganz zu



- Die **Aufgaben** im Versuch waren **leicht zu bewältigen**
- Die **Aufgaben** im Versuch waren **repräsentativ** für die Tätigkeit im Cockpit eines Kampfflugzeugs.
- Das **Versuchsszenario** war **repräsentativ** für eine militärische Luftoperation.
- Der **Kräfteinsatz** (bspw. Anzahl Plattformen, Waffen, etc.) war **optimal**.

5. Sonstige Anmerkungen

Anmerkungen zum Versuch