

Mainz Experimental Psychology Reports (MEPR)



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ



Psychologisches Institut
Allgemeine Experimentelle Psychologie
Wallstraße 3
55122 Mainz
Telefon +49 6131 39-39266
Fax +49 6131 39-39268
<https://experimental.psychologie.uni-mainz.de/>

**Heiko Hecht, Chiara Oftring, Louis Bixenmann
und Oliver Daum**

Methodenleitfaden zu Workload, Situational Awareness und Stress

Eine Handreichung für die Wahl geeigneter Messverfahren für Human-Factors-Fragestellungen

HEIKO HECHT¹, CHIARA OFTRING¹, LOUIS BIXENMANN²
& OLIVER DAUM²

¹Psychologisches Institut - Allgemeine Experimentelle Psychologie,
Johannes Gutenberg-Universität, Mainz

²Zentrum für Luft- und Raumfahrtmedizin der Luftwaffe, Manching

Diesen Artikel zitieren als: Hecht, H., Oftring, C., Bixenmann, L. & Daum, O. (2023). *Methodenleitfaden zu Workload, Situational Awareness und Stress*. (Mainz Experimental Psychology Reports Nr. 7). Mainz: Universität Mainz, Abteilung Allgemeine Experimentelle Psychologie.

<https://experimental.psychologie.uni-mainz.de/forschungsberichte/>

Inhalt

1	Einleitung	4
1.1	Workload	5
1.2	Situational Awareness	7
1.3	Stress	8
1.4	Zusammenhang zwischen Workload, Situational Awareness und Stress.....	11
2	Testverfahren zur Erfassung von Workload	13
2.1	Workload-Skalen	13
2.1.1	Activation Scale.....	13
2.1.2	ATWIT.....	14
2.1.3	DRAWS.....	15
2.1.4	DALI.....	16
2.1.5	ISA.....	18
2.1.6	IWS.....	20
2.1.7	MACE.....	22
2.1.8	MCH.....	24
2.1.9	MD.....	26
2.1.10	NASA-TLX.....	28
2.1.11	OW.....	31
2.1.12	RSME.....	32
2.1.13	SOCC.....	34
2.1.14	SWAT.....	35
2.1.15	SWORD.....	37
2.1.16	VACP bzw. McCracken-Aldrich-Approach.....	39
2.1.17	VOSO.....	42
2.1.18	WCI/TE.....	44
2.1.19	WP.....	46
2.2	Planungshilfen und Empfehlungen zu WL-Erfassung	48
3	Testverfahren zur Erfassung von Situational Awareness	50
3.1.1	CARS.....	50
3.1.2	CLSA.....	52
3.1.3	Cranfield SAS.....	54
3.1.4	MARS.....	56
3.1.5	QASA.....	57
3.1.6	SAGAT.....	59
3.1.7	SALIENT.....	62
3.1.8	SALSA.....	65
3.1.9	SARS.....	66

3.1.10	SART und 3-D SART	69
3.1.11	SASRF	71
3.1.12	SA-SWORD	73
3.1.13	SAVANT	75
3.1.14	SPAM.....	77
3.1.15	Temporal Awareness	79
3.1.16	VESARS.....	80
3.2	Planungshilfen und Empfehlungen zu SA-Erfassung	82
4	Testverfahren zur Erfassung von Stress	84
4.1	Bewertungsprozesse/ Stresswahrnehmung.....	84
4.1.1	ABF.....	84
4.1.2	TICS	87
4.1.3	AVEM	89
4.1.4	IR.....	90
4.1.5	PSS.....	92
4.1.6	Mental Health Fragebogen	94
4.2	Stressbewältigung	96
4.2.1	RESTQ/EBF	96
4.2.2	CISS.....	97
4.2.3	UBV	99
4.2.4	FERUS	101
4.2.5	SVF	102
4.2.6	SCS/D.....	103
4.3	Stresserleben	105
4.3.1	BELA	105
4.3.2	SRS.....	106
4.3.3	PSRS	108
4.4	Physiologische Indikatoren psychischer Bewertung.....	110
4.4.1	Kardiovaskuläre Aktivität	110
4.4.2	Hautleitfähigkeits-niveau und -reaktion (SCL, SCR)	110
4.4.3	Cortisolspiegel.....	111
4.5	Personale Ressourcen.....	112
4.5.1	SWE.....	112
4.5.2	GSE-3.....	113
4.5.3	SSA	114
4.5.4	Pearlin Mastery Scale.....	116
4.5.5	IPC.....	117
4.5.6	IE-4-Skala.....	118

4.5.7	ASF-E	119
4.5.8	SCS.....	120
4.5.9	H-Skalen	122
4.5.10	RS	124
4.5.11	SOC-Skala	126
4.6	Planungshilfen und Empfehlungen zu Stress-Erfassung	127
5	Literatur.....	128
6	Abbildungsverzeichnis.....	139
7	Abkürzungsverzeichnis.....	140

1 Einleitung

Das vorliegende Dokument soll eine Übersicht über Testverfahren aus den drei konzeptuell verwandten Bereichen Workload, Situational Awareness und Stress bieten. Viele Autoren haben sich dieser Aufgabe – meist auf Teilaspekte beschränkt – bereits in den vergangenen Jahren und Jahrzehnten angenommen und umfangreiche Reviews oder Bücher innerhalb dieses Themenbereichs verfasst. Wir haben versucht, die aktuellen Informationen und den aktuellen Validierungsstatus der genannten Verfahren zusammenzutragen und auf die relevantesten Informationen zu reduzieren, ohne dabei den wesentlichen Informationsgehalt zu verlieren. Diese Handreichung soll demnach insbesondere eine Übersicht der Literatur und Testmethodik bieten. Sie soll die Entscheidung für oder gegen das eine oder andere Verfahren erleichtern. Alle, die sich darüber hinaus eingehender mit den Verfahren befassen wollen, finden zu den jeweiligen Verfahren sortiert nach Fragestellung Literaturangaben, die in der Regel zugänglich sind. Aufgrund des Arbeitsfelds mit Schwerpunkten in Luftfahrt, Industrie und Militär sind manche Erkenntnisse zu Testverfahren nicht in wissenschaftlichen Magazinen, sondern häufig als Report der jeweiligen Arbeitseinheit herausgegeben worden. Auf manche Quellen hatten wir daher selbst keinen Zugriff, in anderen Fällen wurden in der Literatur Quellen genannt, die nicht veröffentlicht sind.

Unter anderem durch diesen Umstand stellte die Recherche relevanter Literatur teilweise eine große Herausforderung dar. Bei dieser und bei weiteren Aufgaben wurden wir tatkräftig von Sarah Forst, Anna L. Maier, Ilaria Siriner, Carolin Sweeney und Lea Thomas unterstützt, denen für ihre gewissenhafte Arbeit unser Dank gilt.

Wir skizzieren im Folgenden kurz die drei Themenbereiche Workload, Situational Awareness und Stress dergestalt, dass einerseits klar wird, welche Herausforderungen die Messung der jeweiligen Indikatoren mit sich bringt, und andererseits die Logik der Gliederung der sich anschließenden unterschiedlichen Messverfahren zu Tage tritt.

1.1 Workload

Das Konzept der mentalen Belastung (engl. *mental workload*, hier als WL abgekürzt) ist im Sinne der Arbeitspsychologie oft als die Anforderungen, die auf einen Arbeiter oder eine Arbeiterin einströmen, definiert. Belastung wird so scheinbar klar vom subjektiven Erleben dieser Belastung, nämlich der Beanspruchung, abgegrenzt. Wir verstehen Workload allerdings eher als Überbegriff, der sowohl objektive Belastung als auch subjektive Beanspruchung umfasst. Hiermit nähern wir uns der Auffassung, die in der internationalen Norm [EN ISO 10075](#) durchscheint. Das Konzept der WL wurde bereits in den 1970er Jahren beschrieben (Cain, 2007; Huey & Wickens, 1993), es gibt aber bislang keine einheitliche Definition (Cain, 2007; Hart, 2006; Mitchell, 2000). Für einen aktuellen Überblick gängiger Definitionen von WL und den Versuch, drei Kernbereiche hervorzuheben, Selbstauskunft, Performanz und physiologische Maße, sei auf Longo et al. (2022) verwiesen. Relevant ist WL insbesondere im Kontext von standardisierten und anspruchsvollen Leistungsaufgaben, wie etwa beim Landeanflug eines Flugzeuges, für den mehrere Aufgaben im Cockpit in unterschiedlichen Reihenfolgen oder parallel bearbeitet werden müssen. Ein Ziel in der Erfassung von WL ist die Vermeidung möglicher Überforderungen von Arbeitern durch zu hohe Arbeitsanforderungen (Tattersall & Foord, 1996). Stattdessen soll ein System idealerweise so auf die Arbeiterin (hier im Sinne des Englischen *operator* gemeint) abgestimmt werden, dass sie die Aufgaben optimal bewältigen kann (Hart, 2006; Mitchell, 2000). Hier sehen wir bereits, dass der Kontext, in dem Anforderungen gestellt werden, genauso wichtig sein kann wie der Status der Person, die die Aufgabe zu erfüllen hat – also etwa deren Ausbildung und Grad der Expertise. In anderen Worten befasst sich die Untersuchung von Workload zumeist mit einer Mensch-Maschine- oder Mensch-System-Interaktion und zielt dabei häufig auf Situationsvariablen ab, während sie von einem den Standards entsprechend gut ausgebildeten Arbeiter ausgeht.

Zur Erbringung menschlicher Leistungen sind u. a. Wahrnehmung, Vermittlungsprozesse (logisches Denken, Entscheiden und Urteilen), Kommunikation und motorische Handlungen (vgl. Wierwille et al., 1985) notwendig. Motorische Komponenten sind bspw. im Militär während Bodeneinsätzen wichtig und werden in manchen Testverfahren zusätzlich berücksichtigt (z. B. im VACP Modell von McCracken & Aldrich, 1984). Eine Intention der Messung von WL ist die Vorhersage der Performanz eines Operators anhand der gemessenen Werte. Andererseits wird die Quantifizierung von Workload ebenso genutzt, um zu untersuchen, an welcher Stelle zu hohe kognitive Anforderungen nicht mehr im optimalen

Beanspruchungsbereich liegen. Dies könnte die Überforderung eines Arbeiters bedeuten und möglicherweise zu abfallenden Leistungen (Cain, 2007) und Unfällen führen oder bei dauerhafter Überbelastung auch zu psychischen Erkrankungen (Hart, 2006). Somit kann die genaue Ermittlung der Arbeitsbeanspruchungen insbesondere folgende zwei positive Konsequenzen zeigen (Cain, 2007): (1) Komplexe Systeme können an die systemspezifischen Fähigkeiten und durchschnittlichen Belastungsgrenzen von Arbeitern angepasst werden; (2) die notwendigen Anforderungen an die Arbeiter können mithilfe empirisch untersuchter WL beleuchtet und quantifiziert werden, um jene besser trainieren zu können und kapazität bedingte Engpässe zu vermeiden.

Verfahren zur Erhebung von Workload werden üblicherweise in drei Kategorien eingeteilt: (1) leistungsbasierte Messungen, (2) subjektive Verfahren bzw. Selbsteinschätzungen über die Höhe der Workload, und (3) psychophysiologische Messungen (Cain, 2007; Eggemeier et al., 1991; Tattersall & Foord, 1996). Letztere werden im vorliegenden Dokument nur marginal behandelt, da sie bereits in einem separaten Dokument des Zentrums für Luft- und Raumfahrtmedizin der Luftwaffe gut systematisiert vorliegen (Boumann, Cyrol, Biella und Sammito, 2022). Die konzeptionelle Komplexität des jeweiligen Messinstrumentes zur Erfassung von Workload, also ob es sich bspw. um *eine* eindimensionale oder mehrdimensionale Skala, oder gar um *mehrere* Skalen handelt, ist darüber hinaus ein weiterer wichtiger methodischer Parameter des jeweiligen Verfahrens.

Objektive leistungsbasierte Verfahren lassen sich unterteilen in Fremdbewertungen durch gut ausgebildete Beobachterinnen (bspw. Human Factors Experten oder Pilotenausbilder) und Leistungsmessungen (*performance measures*). Manche Verfahren kombinieren Selbst- und Fremdratings, die separat mit der gleichen Erhebungsmethode erfasst werden (z. B. Cranfield SAS oder SARS in Kombination mit SASRF). Leistungsmessungen werden zumeist computer- oder simulatorbasiert erhoben und vergleichen die Leistung bspw. eines Piloten mit der idealen Flugleistung. Im Folgenden sind mehrere Verfahren dargestellt, die unterschiedliche Facetten von WL abbilden und sich ihrer Messung aus unterschiedlichen konzeptionellen Richtungen nähern.

Als mögliches Problem von **subjektiven Testverfahren** zur Erfassung von WL – und auch zur Erfassung von Situational Awareness – ist dessen Intrusivität im Hinblick auf die zeitgleich ausgeführte Primäraufgabe zu nennen (Tattersall & Foord, 1996). Allerdings konnte bislang nur in wenigen Studien ein deutlicher negativer Effekt der Messung an sich auf die Leistung

gefunden werden, geringe intrusive Effekte wurden hingegen etwa bei den Verfahren ISA (Tattersall & Foord, 1996), Cranfield SAS (Jeannot et al., 2003), QASA (Salmon et al., 2006) und SAGAT (Endsley et al., 1998) festgestellt. Subjektiv die Primäraufgabe begleitende Verfahren werden dennoch gerne genutzt, da sie meist einfach und schnell in der Durchführung sind (J. T. Rehmann et al., 1983; Tattersall & Foord, 1996), eine hohe Augenscheinvalidität mit sich bringen und direkte Einblicke in die subjektiv erlebte WL ermöglichen (Reid & Nygren, 1988), die von objektiven Verfahren nicht erfasst werden können. Die eben beschriebenen methodischen Vor- und Nachteile gelten gleichermaßen für Verfahren, die die sogenannte Situational Awareness erfassen sollen.

1.2 Situational Awareness

Situational Awareness¹ (SA), das Bewusstsein über den Zustand des technischen Systems, mit dem man es gerade zu tun hat, gilt als Voraussetzung für eine gute Entscheidungsfindung in komplexen und dynamischen Systemen (Endsley et al., 1998). Ursprünglich im Kontext der Entscheidungsfindung von Piloten entwickelt, wird das Konstrukt Situational Awareness inzwischen in vielen weiteren Bereichen untersucht. Diese umfassen neben der Luft- und Raumfahrt z. B. Aufgabenmanagement im Nuklearenergiesektor, Autofahren, medizinische Eingriffe u. v. m. (Endsley et al., 1998). Als gängigste Definition von SA soll die Beschreibung von Mica Endsley herangezogen werden, in der SA 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, 1988, S. 97). SA wird nach Endsley demnach als dreischrittige kognitive Leistung verstanden, welche (1) die Wahrnehmung raum- und zeitabhängiger Umweltfaktoren, (2) das Verständnis des Bedeutungsgehalts dieser Faktoren für die aktuelle Situation und (3) die Projektion der Bedeutung der aktuellen Faktoren auf eine zukünftige Situation beinhaltet. Besonders in hochkomplexen automatisierten Systemen, in denen zahlreiche und unterschiedlich relevante situative Faktoren vorliegen können, hat deren adäquate Berücksichtigung und Integration einen hohen Stellenwert (Willems & Heiney, 2002). Je besser einem Operator die Enkodierung und Integration relevanter Situationsfaktoren eines Systems gelingt, desto wahrscheinlicher sind positive Outcomes und desto geringer ist das Risiko von Fehlern (Muñiz et al., 1998). Gelingt dies nicht, wird die SA beeinträchtigt und das System

¹ Wir benutzen den englischen Ausdruck samt Akronym SA, welches sich international eingebürgert hat, obwohl hier das deutsche Äquivalent Situationsbewusstsein existiert, aber mit SA merkwürdig abgekürzt wäre.

wird fehleranfälliger. Um die situative Informationsverarbeitung von Operatoren in einem komplexen dynamischen System zu untersuchen und somit dessen Sicherheit zu evaluieren, sind die Entwicklung, Anwendung und Bewertung von Testverfahren zur Erfassung von SA essenziell (Endsley et al., 1998). Ein weiterer Nutzen der Erfassung von SA ist die Möglichkeit, mit diesem Wissen ein individuelles Training pro Teilnehmer oder Teilnehmerin zu erstellen, welches zur Erhöhung ihrer bzw. seiner SA beiträgt (Muñiz et al., 1998).

Dabei darf nicht außer Acht gelassen werden, dass SA sicherlich von den drei Konzepten dasjenige ist, was am schwersten greifbar ist, was nur teilweise von der Tatsache wettgemacht wird, dass sich mit dem SAGAT (s. Abschnitt 3.1.6) ein Messinstrument durchgesetzt hat und dadurch vereinheitlichend wirkt. Eine der großen Herausforderungen bei der Messung von SA liegt darin, dass es ständig wechselnder Ausprägung unterliegt und Messungen immer nur punktuell sein können. Die Volatilität von SA haben kürzlich im Kontext des autonomen Fahrens Röckel und Hecht (2023) zeigen können, die konstatieren, dass das Aufrechterhalten von Situationsbewusstsein auf übergeordneter Ebene (SA-maintenance) sehr schwer erzielbar ist, wenn es bei einer intensiven Zweitaufgabe nicht ganz unmöglich wird.

1.3 Stress

Stress kann als individuelle Antwort auf extreme Workload bis hin zur Überbeanspruchung aufgefasst werden. Wir haben die Erhebungsmethoden von Stressparametern auf der Basis des transaktionalen Stressmodells (Lazarus & Folkman, 1987) entsprechend ihrer inhaltlichen Ausrichtung entlang des Stressprozesses klassifiziert, z.B. Stresswahrnehmung vs. Stressbewältigung. Dementsprechend wurden auch Fragebögen zur Erfassung personaler Ressourcen als protektive Faktoren (Selbstwirksamkeit, Kontrollüberzeugung, Optimismus, Resilienz und Kohärenzerleben) mitaufgenommen. Diese beeinflussen wesentlich den primären Bewertungsprozess des Stressors. Im Rahmen des Stresserlebens und der Stressreaktivität wird zwischen physiologischen Indikatoren psychischer Aktivierung (objektives Stresslevel) und subjektivem Empfinden (subjektives Stresslevel) unterschieden.

Nach Canon (1914, zit. nach Lazarus & Folkman, 1984) entsteht Stress mit der Wahrnehmung emotional erregender Reize (z. B. Hinweise auf eine akute physische Bedrohung), die durch einen adaptiven Prozess der Wiederherstellung des inneren Gleichgewichtes durch vermehrte Ausschüttung von Adrenalin und deren Folgen charakterisiert ist (emergency reaction, fight or flight reaction). Hans Selye erweiterte den Ansatz von Canon und definiert

Stress allgemein als eine unspezifische Reaktion des Körpers auf eine Anforderung (Lizcke & Schuh, 2010; Selye, 1959). Dies umfasst Reaktionen sowohl auf negative als auch auf positive Erlebnisse und durchläuft drei Phasen – Alarmphase, Widerstandsphase und Erschöpfungsphase. Körperliche sowie psychische Reize können mögliche Stressoren sein, z.B. Infektionen, Verletzungen sowie Probleme in der Partnerschaft oder arbeitsbezogene Überforderung oder Überlastung (zur modernen Einordnung siehe auch Gesundheitsberichterstattung des Bundes, 2022). Unabhängig von der Art des einwirkenden Stressors reagiert der Organismus zunächst mit einer körperlichen Anpassungsreaktion in Form einer vermehrten Ausschüttung von Cortisol, Adrenalin und Noradrenalin sowie einer Erhöhung des Herzschlags und des Blutdrucks. Laut Selye ist diese Anpassungsleistung des Organismus mit einer Schwächung des Immunsystems und somit mit einer verringerten Widerstandsfähigkeit gegenüber anderen Stressoren verbunden. Chronischer Stress kann zu psychischen und neurobiologischen Veränderungen führen, beispielsweise zu einem Volumenverlust im Hippocampus (Wolf, 2014).

Das Transaktionale Stressmodell von Lazarus (Lazarus & Folkman 1987) liefert einen konzeptionellen Erklärungsansatz für die Entstehung von Stress, der den Begriff „transaktional“ als eine dynamische Beziehung zwischen Person und Umwelt auffasst (relationales Konzept). Auf Basis kognitiver (primärer) Bewertungsprozesse werden Reize (z. B. Ereignisse) in Bezug auf ihre subjektive Bedeutung (Qualität) und ihr Ausmaß (Quantität) eingeschätzt. Dementsprechend sind Ereignisse individuums- und erfahrungsabhängig und können als (1) irrelevant, (2) positiv oder (3) negativ und somit stressend bewertet werden. Dabei kann die stressbezogene Bewertung nochmals nach den Beziehungen (1) Schaden-Verlust, (2) Bedrohung oder (3) Herausforderung differenziert werden. Der sekundäre Bewertungsprozess des transaktionalen Stressmodells bezieht sich auf die interaktive Bewertung des Stressors mit den subjektiven Bewältigungsressourcen und -möglichkeiten. Übersteigt die Stressbelastung auf der einen Seite die entgegenstehenden Ressourcen auf der anderen Seite kommt es zu gesundheitlichen Problemen. Im Rahmen der Stressbewältigung werden drei übergeordnete Copingstrategien unterschieden: Problemorientiertes (regulierendes) Coping (Lösung der Situation, z.B. erfolgreiches Flugmanöver), emotionsorientiertes Coping (Abbau der emotionalen Erregung vor einem Flugeinsatz z. B. durch Atemtechniken, Meditation) und bewertungsorientiertes (instrumentelles) Coping (z. B. wenn nach einem erfolgreichen Flugmanöver kommende Manöver positiv bewertet werden, sodass diese zukünftig lediglich als Herausforderung angesehen werden). Ebenso spielen situative Merkmale eine Rolle, wie der Grad der Beeinflussbarkeit, der Prädizierbarkeit oder der zeitlichen Nähe und

Dauer eines Stressors. Inhaltlich lassen sich psychische und physische Stressoren unterscheiden. So kann ein erhöhtes Stresslevel in leistungsrelevanten Situationen einen optimalen Abruf der individuellen Ressourcen verhindern, sowie langfristig zu gesundheitlichen Problemen und Einschränkungen führen (Staller, 2011; Li et al., 2012). Bryant-Lees et al. (2021) betonen im Rahmen einer Studie mit Drohnenpiloten der U.S. Air Force (RPA) die Wichtigkeit stressreduzierender und resilienzfördernder Maßnahmen, angesichts des erhöhten Risikos dieser Piloten für Burnout. Insbesondere die Transzendente Meditation (TM), Mindfulness Meditation und PMR (*progressive muscle relaxation*) zeigen in dieser Reihenfolge als Stressbewältigungsstrategien in verschiedenen Meta-Analysen stressreduzierende Effekte, z. B. Reduzierung des Blutdrucks und Erhöhung einer positiven Stimmung (Manzoni et al. 2008, Jain et al., 2007).

Der Aspekt der Stressbewältigung (Coping) spielt eine zentrale Rolle im Umgang mit Stressoren und ist auch in der Arbeitswelt entscheidend für die Bewältigung entsprechender Herausforderungen. Schwarzer (2004) unterscheidet vier wesentliche Arten der Bewältigung: (1) das reaktive Coping, als Anstrengung mit einem bereits eingetretenen Ereignis umzugehen, (2) das antizipatorische Coping, das sich auf Bemühungen im Zusammenhang mit einer bevorstehenden Bedrohung richtet, (3) das präventive Coping, welches sich mit unbekanntem Risiken in eher ferner Zukunft beschäftigt und (4) das proaktive Coping, das auf eine noch weitere Perspektive ausgerichtet ist. Während sich das reaktive Coping auf bereits vergangene Ereignisse richtet und daher mit einem hohen Grad der Gewissheit verbunden ist, nimmt die Gewissheit der drei anderen Copingformen mit dem Ausmaß ihrer zeitlichen Orientierung an der Zukunft immer weiter ab.

Im Rahmen der Theorie der Ressourcenerhaltung wird Stress mit dem Verlust an Ressourcen bzw. mangelndem Ressourcenerwerb nach entsprechender Verausgabung gleichgesetzt (Hobfoll & Buchwald, 2004). Grundsätzlich basiert die Theorie auf dem grundlegenden Bedürfnis von Menschen, Ressourcen zu erhalten bzw. zu erweitern (Buchwald & Hobfoll, 2013). Dementsprechend ist bereits das Wissen um potentiellen Stress eine Motivationsquelle für den Einsatz proaktiver Bewältigungsstrategien (Buchwald & Hobfoll, 2013). Um einem möglichen Ressourcenverlust vorzubeugen, sowie das Bedürfnis eines Ressourcengewinns zu befriedigen, werden bestehende Ressourcen in zukünftige Ressourcen investiert (Hobfoll & Buchwald, 2004). Das mit der Ressourcenerhaltungstheorie assoziierte multiaxiale Copingmodell integriert individuelle Bewältigungsstrategien im sozialen Kontext. Diesbezüglich ergeben sich neben der Aktiv-Passiv-Achse, die Prosozial-Antisozial-Achse sowie

die Direkt-Indirekt Achse (Hobfoll & Buchwald, 2004). Die Aktiv-Passiv-Achse bildet das Ausmaß der Initiative ab, welche eine Person zur Bewältigung des Problems ergreift (Aktivität bis Vermeidungsverhalten). Dies beinhaltet neben der Bewältigungsaktivität in der Problemsituation auch präventive Schritte (z.B. mentales Training bestimmter Flugmanöver, um das Arbeitsgedächtnis während des Flugmanövers zu entlasten). Prosoziale Copingstrategien beziehen sich auf soziale, wechselseitige Unterstützungsleistungen bei Problemen, während antisoziales Verhalten darauf abzielt, durch verletzendes Verhalten eigene Gewinne zu maximieren. Indirekte Copingstrategien beziehen sich auf strategische, diplomatische Beeinflussungen im sozialen Kontext (Hobfoll & Buchwald, 2004).

1.4 Zusammenhang zwischen Workload, Situational Awareness und Stress

Die drei Konzepte Workload, Situational Awareness und Stress sind stets dann eng miteinander verwandt, wenn die Arbeiterin (*Operator*) anspruchsvolle Aufgaben in einer hochtechnisierten Umgebung ausführt, die sie an ihre Leistungsgrenzen führt. Erfordert die Situation, dass der *Operator* zu lang zu nah an dieser Grenze arbeiten muss, oder sie sogar überschreitet, ist in hohem Grad der Human-Factors-Spezialist gefragt, um Lösungen zu finden, wie die Beanspruchung wieder in einen tolerablen Bereich geführt werden kann. Dazu reicht die Messung von Workload und Stress, sei es mit subjektiven oder objektiven Maßen, nicht aus, selbst wenn moderne integrative Konzepte wie das der kognitiven Anstrengung zugrunde gelegt werden (Oberfeld, 2023). Maße von WL und Stress müssen um die Messung von SA ergänzt werden, womit in ersteren nicht erfasste qualitative Aspekte erschlossen werden. Der Verlust von SA ist ein Indikator für Kritikalität, der nicht auf Workload reduziert werden kann.

Eine hohe SA kann dazu führen, dass Arbeiten unter sehr hohen Belastungen (hohe WL) souverän und fehlerfrei durchgeführt werden können, während dies bei geringerer SA unwahrscheinlicher wird. Häufig wird der Aufmerksamkeitsfokus unter hoher WL derart eingeschränkt, dass nur noch die relevantesten Informationen bewusst verarbeitet werden können (Endsley, 1988) und weniger relevante Hinweise übersehen werden. Solche Wahrnehmungsfehler oder eine fehlerhafte Interpretation von Informationen können zu verheerenden Folgen führen, wie sie von Badke-Schaub et al. (2008) anschaulich beschrieben

wurden. Sie werden oft – mit fragwürdigen Implikationen – als menschliches Versagen bezeichnet. So wird etwa in den Augen der Öffentlichkeit, ein Bild des Piloten vermittelt, der auch unter extremen Bedingungen und hohem Stress stets die richtigen Entscheidungen trifft. Ohne Ansehung seiner Beanspruchung oder systemischer Faktoren, die verhindert haben, dass eine adäquate SA aufrechterhalten werden konnte, wird oft das „menschliche Versagen“ nicht weiter analysiert. Aufgrund der natürlichen Limitation kognitiver Fähigkeiten müssen komplexe Systeme und Teams jedoch hinsichtlich der Situational Awareness und der Workload gründlich analysiert und dann entsprechend optimiert werden. Die vorliegende Handreichung soll den Forscher wie die Praktikerin bei dieser Arbeit unterstützen, soll helfen, sie bestmöglich auszubilden und im fortwährenden Training als Nachschlagewerk dienen. Zur Einordnung der Tests haben wir knappe aber übersichtliche theoretische Überlegungen angestellt, die etwas Ordnung in die große Zahl der existierenden Testverfahren bringen mögen, welche wir im Folgenden der Reihe nach für WL, SA und Stress zusammengefasst darstellen.

2 Testverfahren zur Erfassung von Workload

2.1 Workload-Skalen

2.1.1 Activation Scale

Autoren	Bartenwerfer (1969)
Verfahren	Unidimensional, subjektiv
Datenerhebung	Selbsteinschätzung nach Aufgabe, Paper & Pencil, Einzeltest
Durchführung	Ähnlich zur Rating Scale Mental Effort (RSME, s. unten). Auf einer Skala von 0 bis 270 erfolgt eine Angabe darüber, wie hoch die mentale Aktivierung ist. Referenzpunkte wie „Zeitung lesen“ sind genannt.
Vorbereitung	Vorbereitung einer Aufgabe, die durchgeführt werden soll.
Validierung	<p>Normstichprobe von $N = 557$ (Bartenwerfer, 1969), allerdings ohne Angabe von Kennwerten (Udris, 1976).</p> <p>Untersuchung des Konzepts „Aktivierung“, das über die Activation Scale erhoben wird, in Zusammenhang mit subjektiven Beanspruchungsmerkmalen von Udris (1976) bei $N = 546$. Gefunden wurden signifikante Korrelationen zwischen hoher Aktivierung und Anspannungs- und Ermüdungsaspekten. Allerdings ohne Angabe von Werten. Diese Ergebnisse weisen darauf hin, dass die Activation Scale wahrscheinlich nur Teilaspekte von Beanspruchung abbildet.</p> <p>Die Activation Scale ist sensitiv für verringerte Aktiviertheit bei Fahrern (nach medikamentöser Sedierung) und somit laut Waard (1996, S. 63–64) ein valides Instrument zur Erfassung von Aktiviertheit.</p>
Bewertung	Keine genauen numerischen Angaben zu Diagnostizität und Sensitivität (Waard, 1996), von daher ist die Güte des Verfahrens schwer zu bewerten. Allerdings ist es einfach in der Anwendung und Auswertung.
Verfügbarkeit	Miller (2001)

2.1.2 ATWIT

	Air Traffic Workload Input Technique
Autoren	Stein (1985)
Verfahren	Unidimensional, subjektiv
Datenerhebung	eine Skala, während einer Simulation / Aufgabe in der Flugsicherung
Durchführung	Konzeptuell der ISA ähnlich, untersucht WL speziell bei ATCOs (<i>Air Traffic Control Operator</i>) während einer Simulation oder verschiedenen Operationen, indem die Höhe der aktuellen WL auf einer Skala angegeben werden soll (Tobaruela Arnedo, 2015). Die Skala reicht von 1 (<i>niedrig</i>) bis 10 (<i>hoch</i>) und wird über zehn Knöpfe abgebildet, von denen in regelmäßigen Abständen, während der Simulation einer in der entsprechenden Höhe betätigt werden soll.
Vorbereitung	Simulationen bzw. Aufgaben vorbereiten. Experten auswählen und für die jeweiligen Aufgaben schulen. Vor den eigentlichen Testungen wurden bei Stein (1985) zusätzlich Trainings durchgeführt.
Validierung	Bei Stein (1985) wurden zusätzlich zu den subjektiven WL-Angaben der ATCOs zwei Experten beauftragt, die Workload objektiv einzuschätzen: <ul style="list-style-type: none"> • Korrelationen zwischen objektiver Taskload und über ATWIT erhobener Workload ($r = .79$) • Effekt des Faktors Taskload auf die subjektive Workload gefunden • Hohe Interrater-Reliabilität ($r = .91$), hohe Korrelation der Beobachterschätzungen von WL mit der tatsächlichen Taskload ($r = .91$), und mit den ATWIT Antworten ($r = .86$)
Bewertung	Eine der am häufigsten verwendeten Leistungsmessungen (Tobaruela Arnedo, 2015). Es liegen allerdings wenige Studien zur Anwendung der ATWIT vor, Validierungsstudien konnten mit Ausnahme von Stein (1985) nicht gefunden werden.
Verfügbarkeit	Materialien in Stein (1985) im Anhang abgedruckt

2.1.3 DRAWS

	Defence Research Agency Workload Scale
Autoren	Farmer, Belyavin, et al. (1995); Farmer, Jordan, et al. (1995)
Verfahren	Multidimensional, subjektiv
Datenerhebung	4 Skalen, Paper & Pencil, abhängig von Aufgabentyp computerbasierte Durchführung, Dauer abhängig von Primäraufgaben, Einzel- & Gruppentestung möglich
Durchführung	Es sollen Ratings abgegeben werden zu den vier Skalen (1) Input Demand, (2) Central Demand und (3) Output Demand einer jeweiligen Aufgabe und zum (4) Zeitdruck (time pressure) während der Bearbeitung einer Aufgabe. Die Skala reicht von 0 (<i>no load</i>) bis 100 (<i>fully loaded</i>), kann aber 100 auch übersteigen.
Vorbereitung	Repräsentative Primäraufgaben müssen gewählt oder erstellt werden.
Validierung	<p>Validierungsstudie ($N = 12$) von Newman et al. (1996) über die drei Verfahren RMSE, DRAWS und Bedford Skala anhand von Flugsimulationen als Primäraufgabe.</p> <ul style="list-style-type: none"> <input type="checkbox"/> Zusammenhänge bestanden zwischen den vier DRAWS-Skalen und vier gefundenen Faktoren von WL (aus einer Faktorenanalyse) <input type="checkbox"/> Sensitivität von DRAWS im Vergleich der drei Verfahren am höchsten: Erfasste Schwankungen in Aufgabenschwierigkeit und Erfahrung von Probanden am sensitivsten. <input type="checkbox"/> Konstrukt- und Kriteriumsvalidität (konkurrente Validität) gefunden.
Bewertung	<p>Geeignet für Workload Messung einzelner Aufgaben oder für den Einsatz mit mehreren simultanen Aufgaben. Es handelt sich laut Newman et al. (1996) um ein valides und sensitives Verfahren, allerdings war die Stichprobe über 12 Personen relativ klein, um generalisierbare Aussagen über die Güte des Verfahrens zu treffen.</p> <p>Es wurde außerdem eine prädiktive Form der DRAWS entwickelt, die POP (<i>Prediction of Operator Performance</i>) (ebenfalls Farmer, Belyavin, et al., 1995; Farmer, Jordan, et al., 1995), die eine gute prädiktive Validität für die Vorhersage der Leistung einer Person besitzt (Cain, 2007).</p>

2.1.4 DALI

	Driver Activity Load Index
Autoren	Pauzié und Pachiaudi (1997)
Verfahren	Multidimensional, subjektiv
Datenerhebung	Paper & Pencil, 6 Skalen, Einzel- oder Gruppentestung, post trial, ca. 5 min
Durchführung	<p>Revidierte Version des NASA-TLX, speziell zur Erfassung von Workload während des Autofahrens entwickelt.</p> <p>Sechs Dimensionen (Pauzié, 2008):</p> <ul style="list-style-type: none"> (1) effort of attention (2) visual demand (3) auditory demand (4) temporal demand (5) interference (6) situational stress <p>Diese werden von 0 (<i>low</i>) bis 100 (<i>high mental demands</i>) geratet. Die Auswertung erfolgt wie für den NASA-TLX über Gewichtung der einzelnen Skalen.</p>
Vorbereitung	Fahrsimulation bzw. -szenario muss konzipiert werden
Validierung	(Iranische) Validierungsstudie von Zakerian et al. (2018) mit $N = 100$ Busfahrern. Cronbach's Alpha $> .70$, keine numerischen Angaben zur Validität lesbar, aber laut Autoren sei der DALI reliabel und valide.
Bewertung	Vor dem Hintergrund, dass die Grundlage für den DALI der NASA-TLX ist und der DALI Index speziell für Workload während des Autofahrens entwickelt wurde, scheint das Verfahren für diesen Kontext gut geeignet zu sein. Es liegt allerdings nur eine Studie (auf Persisch) vor, welche den DALI hinsichtlich Reliabilität und Validität untersucht hat (Zakerian et al., 2018).
Verfügbarkeit	Skalen in Pauzié (2008)

Title	Endpoints	Description
effort of attention	low/high	to evaluate the attention required by the activity – to think about, to decide, to choose, to look for and so on
visual demand	low/high	to evaluate the visual demand necessary for the activity
auditory demand	low/high	to evaluate the auditory demand necessary for the activity
temporal demand	low/high	to evaluate the specific constraint owing to timing demand when running the activity
interference	low/high	to evaluate the possible disturbance when running the driving activity simultaneously with any other supplementary task such as phoning, using systems or radio and so on
situational stress	low/high	to evaluate the level of constraints/stress while conducting the activity such as fatigue, insecure feeling, irritation, discouragement and so on

Abbildung 1

Übersicht der Faktoren des DALI und deren Beschreibung. Abbildung aus Pauzié (2008).

2.1.5 ISA

	Instantaneous Self-Assessment
Autoren	Brennan (1992); Jordan und Brennan (1992); Kirwan et al. (1997)
Verfahren	Unidimensional, subjektiv
Datenerhebung	Paper & Pencil, nach oder während einer Simulation, Einzeltestung, Länge abhängig von den Aufgaben während der Simulation
Durchführung	<p>Insbesondere zur Erfassung von WL bei ATCOs entwickelt.</p> <p>Die subjektive Workload soll in eine von fünf Kategorien eingeordnet werden:</p> <p><i>Excessive</i> (5), <i>high</i> (4), <i>comfortable</i> (3), <i>relaxed</i> (2), und <i>under-utilized</i> (1) (Tattersall & Foord, 1996). Ein visueller Hinweisreiz zeigt während einer Simulation an, wann die Workload von 1 bis 5 geratet werden soll.</p>
Vorbereitung	<ol style="list-style-type: none"> 1. Aufgabenbeschreibung erarbeiten 2. Teilnehmer briefen und Beispiele zur Höhe der WL innerhalb der fünf Kategorien geben, damit sie dadurch anschließend die WL treffend einschätzen können 3. Testdurchgang 4. Aufgabe bzw. Simulation starten 5. In regelmäßigen Abständen (z. B. 2 min) die Höhe der WL erfragen (verbal oder per Hinweisreiz und Eingabe per Knopfdruck) 6. WL-Profil auf Basis der Daten erstellen, typischerweise als Graph, der die Höhe der erfragten WL als einzelne Punkte abbildet
Validierung	<p>Tattersall und Foord (1996) untersuchten die WL von $N = 30$ nicht näher beschriebenen Personen, während einer Tracking Aufgabe. Es zeigten sich hochsignifikante Korrelationen der ISA Ratings mit der Aufgabenleistung ($r = .60$) und der mentalen Anstrengung ($r = .73$).</p> <p>Leggatt (2005) bewertet die Validität auf Basis eines experimentellen Ansatzes als hoch: $N = 30$ TeilnehmerInnen, darunter 15 RAF C30 PilotInnen, und 15 MitarbeiterInnen der RAF (z.B. Ingenieure) führten die Multi-Attribute Task Battery (MATB) durch.</p> <p><input type="checkbox"/> Konkurrente Validität: $r = .55$ mit NASA-TLX</p>

	<ul style="list-style-type: none"> <input type="checkbox"/> Konstruktvalidität: $r = .82$ mit Aufgabenanspruch der MATB <input type="checkbox"/> keine signifikante Korrelation mit der Aufgabenleistung
Bewertung	<p>Ergebnisse des ISA sind eng korreliert mit dem SWAT Test (Tattersall & Foord, 1996) und dem NASA-TLX (Leggatt, 2005). Kritisch ist hier anzumerken, dass die visuelle Abfrage eventuell intrusiv und nachteilig für die währenddessen ausgeführte Hauptaufgabe (z. B. Autofahren) sein kann (Tattersall & Foord, 1996). Leggatt (2005) konnte allerdings keine Performanzunterschiede zwischen Aufgaben mit und ohne ISA finden.</p> <p>Insgesamt ist das ISA eingeschränkt empfehlenswert, es hat gegenüber dem SWAT das Nachsehen und wird häufig um andere Verfahren ergänzt, um die WL dadurch möglichst umfassend abzubilden (Jordan & Brennan, 1992). Es ist insgesamt ein sehr einfaches und ökonomisches Verfahren, das sensitiv und auch valide ist, aber gleichzeitig intrusiv sein kann und Workload nur teilweise abbildet.</p>
Verfügbarkeit	Jordan und Brennan (1992)

Level	Workload Heading	Spare Capacity	Description
5	Excessive	None	Behind on tasks; losing track of the full picture
4	High	Very Little	Non essential tasks suffering. Could not work at this level very long.
3	Comfortable Busy Pace	Some	All tasks well in hand. Busy but stimulating pace. Could keep going continuously at this level.
2	Relaxed	Ample	More than enough time for all tasks. Active on ATC task less than 50% of the time.
1	Under- Utilised	Very Much	Nothing to do. Rather boring.

Abbildung 2
Skala und Beschreibung der ISA aus Kirwan et al. (1997)

2.1.6 IWS

	Integrated Workload Scale
Autoren	Pickup et al. (2005)
Verfahren	Unidimensional, subjektiv
Datenerhebung	<ul style="list-style-type: none"> <input type="checkbox"/> Paper & Pencil: Werte werden von VL mittels Stoppuhr in festgelegten Intervallen abgefragt und der Teilnehmer gibt die Höhe der WL mithilfe einer ausgedruckten, farbkodierten neunstufigen Skala an. <input type="checkbox"/> Oder: Computerbasierte Erhebung mittels speziellen Tastenfeldes, das an einen Laptop angeschlossen ist <input type="checkbox"/> Einzeltestung, Länge der Durchführung abhängig von Länge der Primäraufgaben
Durchführung	<p>Echtzeitmessung von Workload, insb. im Zugverkehr und der -sicherung. Neun-stufige Skala, die von <i>not demanding</i> bis <i>too demanding</i> reicht. Die neun Stufen sind farblich kodiert. Als Vorlage zur Entwicklung der IWS diente das ISA.</p> <p>Vier Dimensionen:</p> <ul style="list-style-type: none"> (1) load (2) demand (3) effort (4) effects <p>Während einer Aufgabe wird in regelmäßigen Zeitabständen (z. B. 2 min) die Höhe der Workload abgefragt und subjektiv durch die untersuchte Person bewertet.</p>
Vorbereitung	Primäraufgaben vorbereiten (z. B. Simulation entwerfen), Erhebungsmodus (analog oder digital) vorbereiten
Validierung	<p>Pickup et al. (2005): Pilotstudie mit Human Factors Experten und Subject Matter Experten (SME), und $N = 130$ Zugsicherungsarbeitern (<i>signalers</i>). Anschließende Validierung in zwei Simulator-Studien:</p> <ul style="list-style-type: none"> - $N = 7$ Signaler, hohe Korrelation mit SWAT: $r = .69$

	- N = 10 Signaler und 1 Ingenieur, Ergebnisse der IWS bildeten tatsächliche Workload ab
Bewertung	Umfangreiche methodologische Überlegungen vor und während der Erstellung, um Gütekriterien zu erfüllen (Pickup et al., 2005). Die IWS erweist sich als ein valides und sensitives Verfahren, um die Workload im Zugverkehr und der Zugsicherung zu erheben. Sie ist wenig intrusiv und die Art der Datenerhebung ist variabel auf das individuelle Erhebungssetting anpassbar. Insgesamt kann die IWS in diesem spezifischen Arbeitsbereich überzeugen. Die Übertragbarkeit auf andere Bereiche ist allerdings fraglich.
Verfügbarkeit	Abbildung der Skala, des Versuchsaufbaus und der Ergebnisse in Pickup et al. (2005).

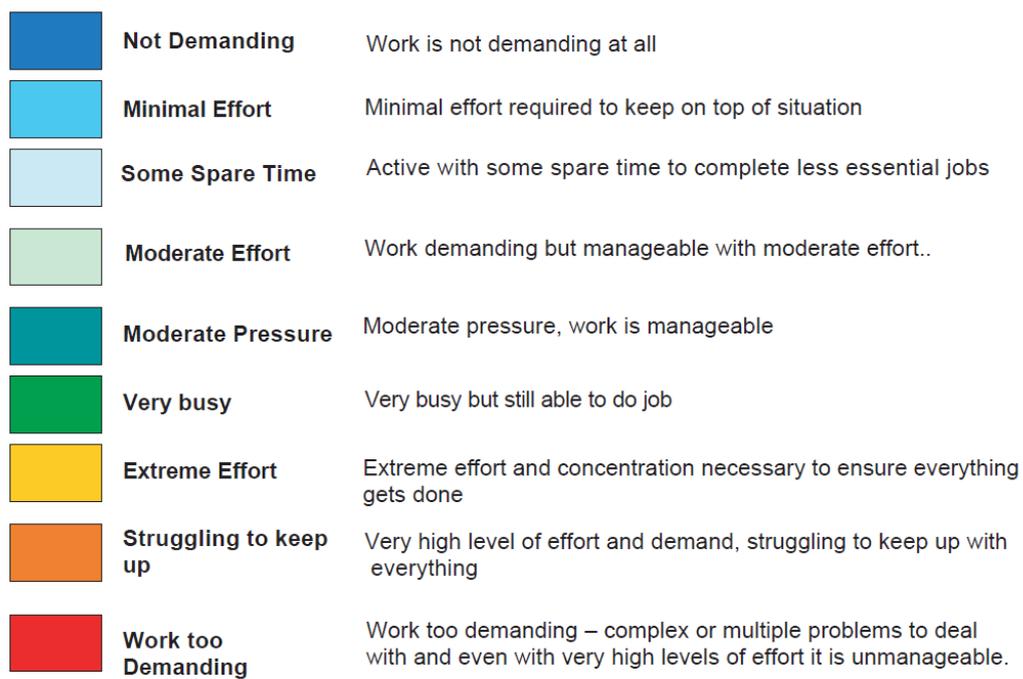


Fig. 2. IWS for rail signallers (colour codes run from blue for not demanding to red for work too demanding).

Abbildung 3

Farblich kodierte Skala der IWS aus Pickup et al. (2005).

2.1.7 MACE

	Malvern Capacity Estimate
Autoren	Goillau und Kelly (1997)
Verfahren	Unidimensional, subjektiv
Datenerhebung	Länge der Erhebung abhängig von Aufgabendauer, ein Messzeitpunkt nimmt wenige Sekunden in Anspruch. Eine Skala. Nach Stanton et al. (2013) meistens digital angewendet. Eher in Einzeltestungen eingesetzt, Gruppenkontext auch denkbar.
Durchführung	Für Piloten entwickelt. Es soll die Frage geratet werden, mit wie viel mehr oder weniger Luftverkehr eine Pilotin im realen Leben zurechtkäme. Die Skala reicht von -100% (<i>less traffic than this run</i>) über 0% (<i>same traffic as this</i>) bis +100% (<i>more traffic than this</i>) (Goillau & Kelly, 1997). Indirekte Messung von Workload über den Vergleich von Kapazität und Workload (Tobaruela Arnedo, 2015).
Vorbereitung	<ol style="list-style-type: none"> 1. Einweisung der Teilnehmer, damit diese ein Verständnis für die Skala und die verschiedenen Stufen (z.B. Bedeutung von 50 %) entwickeln können, Erstellung von Beispielen für jedes Skalenlevel 2. Probedurchlauf mit der Möglichkeit für Fragen 3. Simulation vor dem Rating
Validierung	Laut Goillau und Kelly (1997) sind die WL-Schätzungen sehr konsistent. Weitere Validierungsstudien sind vorhanden, wurden aber nicht publiziert, somit existiert keine zugängliche Literatur zur Validität des MACE (Stanton et al., 2013).
Bewertung	Der MACE wurde teilweise validiert und zeigt eine gute Akzeptanz. Die verfügbaren Kapazitäten werden bei höherer Workload scheinbar besser eingeschätzt, somit besteht kein linearer Zusammenhang (Goillau & Kelly, 1997). Als einfacher Vergleich zwischen aktueller Workload und möglicher Spitzenleistung ist der Einsatz der MACE durchaus sinnvoll. Generell zeigt er die Vorteile unidimensionaler Verfahren (einfach, schnell, ökonomisch und trotzdem aussagekräftig), andererseits sind die erhobenen Informationen

	subjektiv und nicht so umfangreich wie bei multidimensionalen Verfahren. Eine Übertragung auf den Straßenverkehr ist eventuell nicht 1:1 möglich.
Verfügbarkeit	Goillau und Kelly (1997)

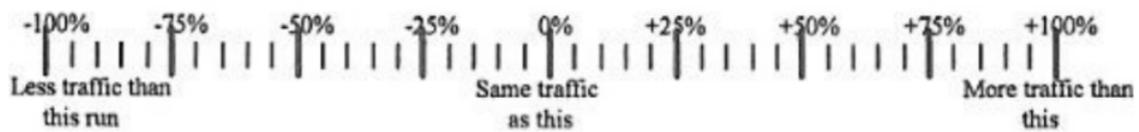


Abbildung 4
Rating-Skala der MACE aus Goillau und Kelly (1997)

2.1.8 MCH

	Modified Cooper Harper Scales
Autoren	Cooper und Harper Jr (1969): Cooper Harper Scale Wierwille und Casali (1983): Modified Cooper Harper Scale
Verfahren	Unidimensional, subjektiv
Datenerhebung	Paper & Pencil, Post Trial, Gruppentestung möglich, wenige Minuten
Durchführung	Die MCH nutzt einen Entscheidungsbaum, um Workload zu erheben. Wurde ursprünglich für die Workload-Messung bei Piloten entwickelt. Die Erhebung erfolgt erst nach dem Trial und nicht währenddessen. Erhoben wird die WL über eine zehnstufige Rating Skala, die in einem globalen WL Rating resultiert. In der ursprünglichen Form (Cooper-Harper Scale) lag der Fokus auf motorischem Arbeiten und Ansprüchen an Arbeiter. In der modifizierten Form (Wierwille & Casali, 1983) werden Workload-Ansprüche speziell auf perzeptueller, kognitiver und der Kommunikations-Ebene gemessen.
Vorbereitung	Schulung über die Nutzung der MCH Skala vor den Trials (Mansikka et al., 2019).
Validierung	Mansikka et al. (2019): Untersuchung der Sensitivität und Validität der MCH bei einer Flugsimulationsaufgabe: <ul style="list-style-type: none"> <input type="checkbox"/> Sensitiv für verschiedene Leistungsniveaus ($p < .01$) <input type="checkbox"/> Konvergente Validität $r = .65$ bis $.90$ mit NASA-TLX über die verschiedenen Leistungsniveaus Hill et al. (1992): Vergleich der MCH Skala mit anderen WL-Verfahren: <ul style="list-style-type: none"> <input type="checkbox"/> Schlechteste Sensitivität im Vergleich mit TLX, OW und SWAT: $r = .80$ mit dem "operator workload factor" (faktorenanalytisch aus den aggregierten Daten zur Aufgabenleistung ermittelt) Wierwille und Casali (1983): Modifikation der <i>Cooper Harper Scale</i> und Validierung <ul style="list-style-type: none"> <input type="checkbox"/> Signifikante Unterscheidung zwischen verschiedenen Belastungsniveaus gefunden ($p < .05$)

	☐ Sensitives Verfahren
Bewertung	Schnelles und einfaches Vorgehen, reliabel und valide. Es sollte aber eine Aufgabenanalyse vor Nutzung der MCH Scales durchgeführt werden. Durch die originäre Anwendung bei Piloten ist die Anwendung der MCH lediglich eingeschränkt trotz vieler Vorteile zu empfehlen.
Verfügbarkeit	Ausführlich beschrieben und abgebildet in Wierwille und Casali (1983) Cooper Harper Scale abgebildet in Miller (2001)

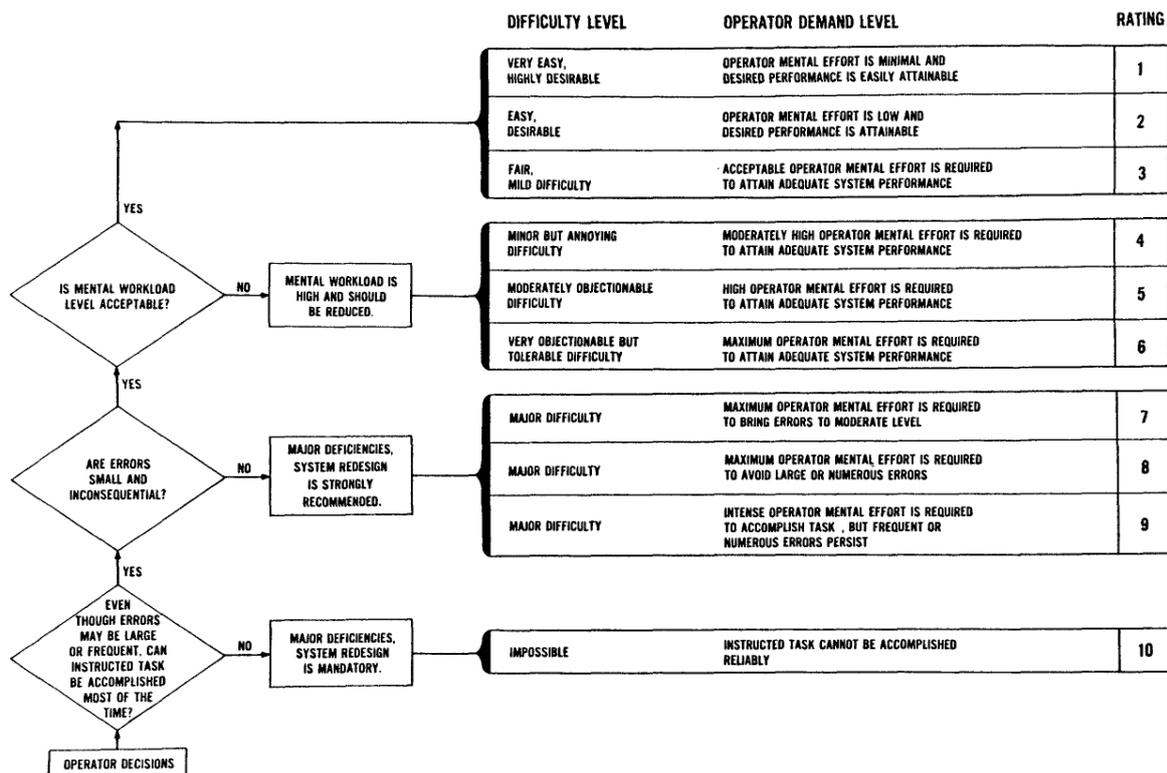


Abbildung 5
Die MCH, aus Wierwille und Casali (1983)

2.1.9 MD

	Multi-descriptor Scale
Autoren	Casali und Wierwille (1983)
Verfahren	Multidimensional, subjektiv
Datenerhebung	Posttrial, Rating direkt nach der Aufgabe, 11-stufige Skala, Einzeltestung, simulator-/computerbasiert
Durchführung	<p>Für die Luft- und Raumfahrt entwickelt.</p> <p>Enthält sieben Bestandteile, welche Workload abbilden sollen:</p> <ol style="list-style-type: none"> (1) attentional demand (2) error level (3) difficulty (4) task complexity (5) mental workload (6) stress level (7) overload level <p>Die Erhebung folgt nach einer Aufgabe, die meistens eine Flugsimulation darstellt. Die sieben Dimensionen sollen auf einer elfstufigen Skala geratet werden. Für jede Aufgabenrunde wird ein Mittelwert je Skala und ein Mittelwert der gesamten WL berechnet.</p>
Vorbereitung	Die Simulatorentaufgaben und ein Übungsflug vor Beginn der eigentlichen Testung müssen erstellt und durchgeführt werden.
Validierung	<p>Casali und Wierwille (1983)</p> <ul style="list-style-type: none"> <input type="checkbox"/> MD ist sensitiv für Veränderungen der Aufgabenbelastung, allerdings sind Unterscheidungen nur zwischen hoher und niedriger Belastung signifikant, nicht zwischen mittelhohen Loads <input type="checkbox"/> Weniger sensitiv als die Modified Cooper-Harper Scale (MCH) <ul style="list-style-type: none"> ➔ Nur ein sinnvolles Verfahren bei großen Belastungsunterschieden zwischen den Bedingungen <p>Wierwille et al. (1985): MD wurde überraschenderweise in einem groß angelegten Vergleich von 16 verschiedenen WL-Verfahren als nicht sensitiv befunden, es konnte lediglich ein Trend ($p = 0.099$) erkannt werden.</p>

Bewertung	<p>Sensitives Verfahren zur Erfassung von Workload, aber weniger sensitiv als die MCH (Casali & Wierwille, 1983). In einer späteren Studie von Wierwille et al. (1985) wurde keine Sensitivität für dieses Verfahren gefunden. Es ist tendenziell gut einsetzbar, insbesondere zur Messung deutlicher Unterschiede in der Workload, aber andere Verfahren können mit besseren Gütekriterien überzeugen.</p> <p>Ggf. können die Ratings auch während der Belastungsphase eingesetzt werden. Kintz et al. (2023) fanden für den Bereich der Raumfahrt heraus, dass eingebettete Systeme nicht-intrusiv während der Aufgabenbearbeitung Informationen erfassen können, die mit WL und auch mit SA zusammenhängen. In Kombination mit weiteren demographischen Informationen und beobachtbaren Informationen über die Aufgabe können diese sowohl WL als auch SA vorhersagen. Ohne die Kombination mit zusätzlichen Prädiktoren sollten diese Maße jedoch nicht eingesetzt werden.</p>
Verfügbarkeit	Beschrieben in Casali und Wierwille (1983), aber nicht abgebildet

2.1.10 NASA-TLX

	NASA Task Load Index
Autoren	Hart und Staveland (1988)
Verfahren	Multidimensional, subjektiv
Datenerhebung	<ul style="list-style-type: none"> <input type="checkbox"/> 6 bipolare Skalen mit jeweils 21 Abstufungen, <input type="checkbox"/> 12 cm lange Linie repräsentiert die Ratingskala, auf der ein diskretes numerisches Rating abgegeben werden soll <input type="checkbox"/> Paper & Pencil und computerbasierte Version vorhanden <input type="checkbox"/> Statistischer Gewichtungprozess zur Auswertung <input type="checkbox"/> Durchführung dauert wenige Minuten, insgesamte Dauer von Aufgabenlänge abhängig
Durchführung	<p>Umfasst sechs bipolare Subskalen:</p> <ol style="list-style-type: none"> (1) Mental demands (2) Physical demands (3) Temporal demands (4) Own performance (5) Effort (6) Frustration <p>Von 0 (<i>low mental demands</i>) bis 100 (<i>high mental demands</i>) sollen in 21 möglichen Schritten die jeweiligen Subskalen geratet werden. Dadurch wird deutlich, in welchem Bereich welche Aufgabe mehr oder weniger Ressourcen bzw. WL benötigt. Alle Subskalen sind genau definiert (s. Abbildung 6).</p> <p>Mit dem NASA <i>Raw</i> TLX (NASA-RTLX) (vgl. Hart, 2006) wurde versucht, die aufwändige Auswertung zu vereinfachen, indem die Summe aller Subskalen durch sechs geteilt wird. Im Kontext von Autofahren erwies sich der RTLX Skala sogar als sensitiver für <i>mental demand</i> und <i>difficulty in driving</i> als die TLX Skala (Park & Cha, 1998, in Hart, 2006; Miller, 2001). Es wurden allerdings auch vergleichsweise geringere und gleichhohe Sensitivitäten des RTLX gefunden (Hart, 2006).</p>

Vorbereitung	Aufwändige Auswertung nach der Erhebung mit der Berechnung von gewichteten WL-Werten pro Dimension muss beherrscht werden, es gibt allerdings auch online Software, die bei der Auswertung unterstützt (Hart, 2006).
Validierung	<p>Rubio et al. (2004) untersuchten die Sensitivität und Validität des NASA-TLX und verglichen diese mit SWAT und WP.</p> <ul style="list-style-type: none"> □ In einer Single Task Bedingung: NASA-TLX sensitiv für die Aufgabenschwierigkeit ($p < 0.001$), operationalisiert durch die Breite eines Pfades in einer Spurhalteaufgabe □ In einer Dual Task Bedingung: sensitiv für die Aufgabenschwierigkeit in beiden Aufgaben ($p = 0.003$ und $p < 0.001$), in dem Fall auch in der Gedächtnisaufgabe □ Konvergente Validität: Hohe Korrelationen mit WL erfasst durch SWAT ($r = .98$) und WP ($r = .99$) □ Konkurrente Validität: Hohe Korrelationen mit der Aufgabenleistung ($r = .75$ mit Reaktionszeit in der Gedächtnisaufgabe und $r = .65$ mit den Fehlern in der Spurhalteaufgabe).
Bewertung	<p>Sensitives Verfahren, das wenig Bearbeitungszeit in Anspruch nimmt, valide und reliabel ist und außerdem durch seine verschiedenen Skalen mehr Informationen zum Workload der jeweiligen Subskalen pro Aufgabe erfasst als bspw. die SWAT (Hart & Staveland, 1988).</p> <p>Geringe interne Konsistenz, weil es sich um ein subjektives Verfahren handelt. Die Auswertung ist statistisch durch die Rechnung mit Gewichten etwas anspruchsvoll (Miller, 2001).</p> <p>Hart (2006) untersuchte 550 Studien, in denen die NASA-TLX Verwendung findet. Sie wurde inzwischen auf fünf Kontinenten eingesetzt, mehrfach übersetzt und an verschiedenste Kontexte angepasst. Für jede Variante des NASA-TLX sollten die Gütekriterien allerdings extra überprüft werden, insbesondere, wenn Subskalen gelöscht, erweitert oder inhaltlich anders übersetzt werden. Wenn diese Anpassungen methodologisch korrekt vorgenommen werden, bleibt der NASA-TLX weiterhin das vorherrschende Instrument zur reliablen, sensitiven und validen Erfassung der WL.</p> <p>→ Neben der SWAT das beste und somit meistgenutzte Verfahren, um Workload zu erfassen</p>

Verfügbarkeit	Hart und Staveland (1988)
---------------	---------------------------

RATING SCALE DEFINITIONS		
Title	Endpoints	Descriptions
MENTAL DEMAND	<i>Low/High</i>	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	<i>Low/High</i>	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	<i>Low/High</i>	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	<i>good/poor</i>	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	<i>Low/High</i>	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	<i>Low/High</i>	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

13

Abbildung 6

Definition der Subskalen des NASA-TLX, aus Hart und Staveland (1988)

2.1.11 OW

	Overall Workload Scale
Autoren	Vidulich und Tsang (1987)
Verfahren	Unidimensional, subjektiv
Datenerhebung	20-stufige Skala, Paper & Pencil oder digital möglich, Einzeltestung und Gruppentestung möglich, kurze Erhebungsdauer
Durchführung	Auf einer bipolaren zwanzigstufigen Skala soll von 0 (<i>sehr niedrig</i>) bis 100 (<i>sehr hoch</i>) angegeben werden, wie hoch die WL ist. Die angegebene Zahl wird zum nächsten Fünferschritt aufgerundet (z. B. wird 22 auf 25 gerundet und 27 auf 30).
Vorbereitung	Im Vergleich zu anderen Verfahren wenig Vorbereitung notwendig.
Validierung	Bei Hill et al. (1992) zeigt die OW eine signifikante Korrelation mit einem durch eine Faktorenanalyse aggregierten <i>operator workload factor</i> . <ul style="list-style-type: none"> <input type="checkbox"/> $r = .87$ bis $r = .93$ je nach betrachteter Studie <input type="checkbox"/> Korrelation im Vergleich mit drei anderen, teils komplexeren Verfahren (TLX, MCH, SWAT) nach TLX am zweithöchsten
Bewertung	Qualität der Ergebnisse mit denen aus dem NASA-TLX vergleichbar (Miller, 2001). In etwa genauso sensitiv wie multidimensionale Skalen (Hill et al., 1992). Laut Hill et al. (1992) ist die OW ein nützliches Screeningtool, um mögliche Engpässe zu administrieren, an denen die WL zu hoch sein könnte.
Verfügbarkeit	Miller (2001)

2.1.12 RSME

	Rating Scale Mental Effort
Autoren	Zijlstra und van Doorn (1985); Zijlstra und Meijman (1989); Zijlstra (1993)
Verfahren	Unidimensional, subjektiv
Datenerhebung	Computerbasiert, post-Trial Erhebung der WL, 15-stufige Skala, Einzeltestung
Durchführung	Auf einer 150 mm langen Skala mit 15 Abstufungen wird der <i>invested effort</i> geratet (Waard, 1996), der von <i>not at all effortful</i> bis <i>very effortful</i> reicht (Newman et al., 1996).
Vorbereitung	Vorbereitung der Primäraufgabe
Validierung	Bei Brookhuis et al. (2009) zeigt sich die RSME Skala sensitiv für unterschiedliche Aufgabenbedingungen: <ul style="list-style-type: none"> <input type="checkbox"/> Höheres WL-Rating bei nebligen Fahrbedingungen ($p < 0.01$), wenn kein Fahrassistent genutzt wurde <input type="checkbox"/> Höheres WL-Rating ohne Fahrassistenten ($p < 0.001$)
Bewertung	Wird eher selten genutzt, ist aber simpel anwendbar, einfach auszuwerten und gut einsetzbar, um eine subjektive Einschätzung des investierten Mental Effort zu erhalten (Waard, 1996). Workload wird hier also ähnlich wie beim MACE nur indirekt erhoben. Das Verfahren kann als sensitiv angesehen werden, weitere Validierungsstudien liegen allerdings nicht vor, weswegen es insgesamt eingeschränkt zu empfehlen ist.
Verfügbarkeit	Miller (2001)

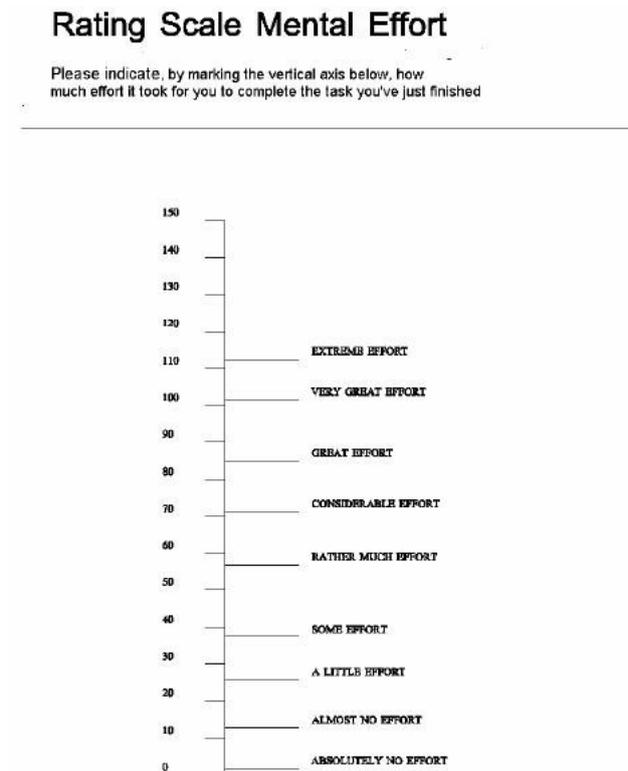


Abbildung 7
Skala der RSME, aus Miller (2001)

2.1.13 SOCC

	Subjective Opinion via Continuous Control
Autoren	Antin und Wierwille (1984)
Verfahren	Unidimensional, subjektiv
Datenerhebung	Kontinuierliche Erfassung der WL über eine Handsteuerung, computerbasiert, Einzeltestung, Dauer von Aufgabe abhängig
Durchführung	Teilnehmer bewegen eine Hand-Steuerung, die sich um ca. 90 Grad rotieren lässt, mit einer ansonsten unbenutzten Hand zu einem Minimum-, Medium- oder Maximum-Punkt. Dies ermöglicht eine momentabhängige Angabe der WL und der WL-Veränderungen (Wierwille & Eggemeier, 1993).
Vorbereitung	
Validierung	<p>Antin und Wierwille (1984) untersuchten die Fähigkeit des Verfahrens, die momentane WL-Belastung zu messen (<i>instantaneous mental workload, IMWL</i>).</p> <ul style="list-style-type: none"> <input type="checkbox"/> Vorhersage des SOCC Ratings über die Aufgabenbelastung einer visuellen Suchaufgabe möglich mit $R^2 = .46$ <input type="checkbox"/> Bei Mittelung über die Teilnehmerinnen hinweg Vorhersage des SOCC Ratings durch Aufgabenbelastung mit $R^2 = .71$ möglich <input type="checkbox"/> Negativ merken die Autoren eine hohe Zwischen-Subjekt Variabilität an
Bewertung	Sehr sensitiv bei kurzen Intervallen von mentaler WL (Miller, 2001). Es ist allerdings problematisch, dass permanent eine Hand für die Nutzung von SOCC in Gebrauch ist. Antin und Wierwille (1984) schlagen daher andere Kontrollmöglichkeiten, wie z. B. eine Fußsteuerung, vor.
Verfügbarkeit	Miller (2001)

2.1.14 SWAT

	Subjective W orkload A ssessment T echnique
Autoren	Reid et al. (1981); Reid und Nygren (1988)
Verfahren	Multidimensional, subjektiv
Datenerhebung	3 Dimensionen, insgesamt 9 Items, dreistufige Ratingskala, Kartensortierungsaufgabe von 27 Kombinationen, anschließend gruppierte statistische Auswertung, Einzel- und Gruppentestung, aufwändiges Verfahren
Durchführung	<p>Die Workload wird über die drei Dimensionen <i>time load</i>, <i>mental effort load</i> und <i>psychological stress load</i> operationalisiert, die jeweils als low (1), medium (2) oder high (3) geratet werden sollen (Pauzié, 2008; Reid & Nygren, 1988). Pro Dimension werden drei Items erfasst, sodass sich ein dreidimensionales Konstrukt von WL mit $3 \text{ Dimensionen} * 3 \text{ Items} * 3 \text{ Ausprägungen} = 27 \text{ Zellen}$ ergibt (s. Abbildung 8). Die Definition der drei Ratingstufen sind für alle drei Dimensionen in Reid und Nygren (1988, S. 199) abgedruckt.</p> <p>Die Durchführung der SWAT ist in zwei Schritte gegliedert: Skalenentwicklung und Bewertung der entwickelten Items mittels einer Kartensortierungsaufgabe (s. Vorbereitung) und anschließendes <i>Event Scoring</i> (s.u.).</p> <p>Nach der Skalenentwicklung und Kartensortierung wird die Übereinstimmung in der Reihenfolge der sortierten Karten zwischen den Teilnehmern mittels Kendalls Konkordanzkoeffizienten W berechnet. Teilnehmerinnen mit hoher Übereinstimmung ($W > .75$) bzgl. der Sortierung werden anschließend in Gruppen zusammengefasst (wenn $W < .75$ siehe Reid & Nygren, 1988, S. 200 f.).</p> <p>Pro Gruppe wird jeder der 27 Zellen ein Wert zwischen 0 und 100 zugeteilt, wobei auf gleiche Abstände geachtet wird, sodass die Höhe der WL jeder Kombination nun über eine metrische Skala abgebildet wird.</p> <p>Neue Situationen oder Ereignisse können nun anhand der Kombinationsmatrix einer Zelle zugeordnet werden und der jeweilige Wert (von 0 bis 100) kann für diese Situationen übernommen werden. Dieses Vorgehen wird von den Autoren <i>Event Scoring</i> genannt.</p>
Vorbereitung	<p>Drei Schritte sind vor Nutzung der SWAT nötig:</p> <ol style="list-style-type: none"> 1. Scale development: alle möglichen Kombinationen ($3*3*3$) werden auf 27 Karten dargestellt.

	<p>2. Pro Karte bzw. Kombination wird die WL von Experten (und anschließend während einer Testung von Teilnehmern) eingeschätzt und aufsteigend sortiert.</p> <p>3. Jeder Kombination wird ein Wert zwischen 0 und 100 zugeteilt</p> <p>Aus den Beurteilungen können Scores für die drei einzelnen Dimensionen sowie ein Gesamt-Workload-Score generiert werden (Additives Modell). Bei Warr et al. (1986) erfolgt außerdem ein Training der SWAT Nutzung vor der eigentlichen Anwendung.</p>
Validierung	<p>Reid und Nygren (1988): Retest-Reliabilität und prädiktive Validität vorhanden. Nach Nygren (1991) ist SWAT (im Vergleich zum TLX) vor allem bei niedriger Workload weniger sensitiv, da die dreistufige Skala (<i>low, medium, high</i>) die Breite der Dimensionen einschränkt und die geringstmögliche Angabe <i>low</i> ist. Es ist also unklar, ob die Bewertung als <i>low</i> ein niedriges oder ein Null-Level meint. Xiao et al. (2005) bewerteten die Validität der SWAT als gut: Konvergente Validität von $r = 0.49$ mit NASA-TLX.</p>
Bewertung	<p>Valide, reliabel und sensitiv, in vielen Reviews als gute Alternative zum NASA-TLX genannt (z. B. Cain, 2007; Rubio et al., 2004), aber aufwändig in der Vorbereitung.</p> <p>Die SWAT wird immer wieder als das beste Alternativverfahren zum NASA-TLX genannt, wenn der Aufwand der Skalenentwicklung bewältigt werden kann, scheint dieses Verfahren eine der besten Möglichkeiten darzustellen.</p>
Verfügbarkeit	<p>Reid und Nygren (1988); Miller (2001)</p>

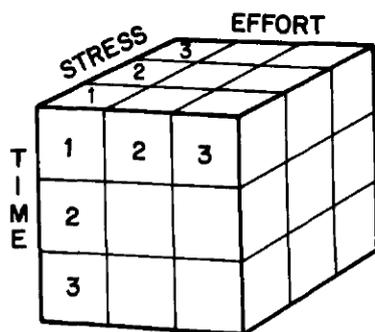


Figure 1. Three-Dimensional Workload Construct

Abbildung 8

Dreidimensionales Workload Konstrukt der SWAT, Abbildung aus Reid und Nygren (1988)

2.1.15 SWORD

	Subjective Workload Dominance
Autoren	Vidulich et al. (1991); Vidulich und Hughes (1991)
Verfahren	Komparatives Verfahren, subjektiv
Datenerhebung	Rating post-Trial, Dauer von Aufgabenmenge abhängig, Einzeltestung, Paper & Pencil
Durchführung	<p>Es handelt sich um ein komparatives Verfahren zwischen mehreren Aufgaben. Die Workload soll für jede Aufgabe im Vergleich zu zwei weiteren Aufgaben beurteilt werden. Die Skala umfasst 17 Schritte. Es wird statistisch ausgewertet, wie konsistent die Urteile in Bezug auf die WL ausfallen (Casner & Gore, 2010).</p> <p>Verschiedene statistische Auswertungsverfahren sind vorhanden (Budescu et al., 1986; Williams & Crawford, 1980).</p>
Vorbereitung	Entwicklung des SWORD Antwortbogens mit allen möglichen Kombinationen zweier Aufgaben. Entwicklung der SWORD Matrix mit den Aufgaben als Zeilen und Spalten.
Validierung	<p>Funke et al. (2013) vergleichen SWORD mit anderen subjektiven Verfahren.</p> <ul style="list-style-type: none"> <input type="checkbox"/> SWORD sensitiv für die Aufgabenschwierigkeit, signifikante Mittelwertsunterschiede zwischen allen drei Schwierigkeitsstufen ($p < .05$) <input type="checkbox"/> Konkurrente Validität: Keine signifikante Korrelation mit der Aufgabenleistung ($r = .24$), aber höher als bei TLX, MRQ und WP <input type="checkbox"/> Konvergente Validität: Keine signifikante Korrelation mit den subjektiven Verfahren TLX, WP oder MRQ
Bewertung	Sinnvoll bei Vergleichen zwischen verschiedenen Aufgaben, wird aber erst nach Ausführung aller Aufgaben erhoben, also im Kontext von bspw. Autofahren eher ungeeignet oder eben anfälliger für Verzerrungen.
Verfügbarkeit	Casner und Gore (2010)

	Absolute	Very Strong	Strong	Weak	EQUAL	Weak	Strong	Very Strong	Absolute
TASK A	---	---	---	---	---	---	---	---	TASK B
TASK A	---	---	---	---	---	---	---	---	TASK C
TASK B	---	---	---	---	---	---	---	---	TASK C

Figure 5. SWORD response sheet.

Abbildung 9

Antwortbogen des SWORD, Abbildung aus Casner und Gore (2010)

MINOR Attitude Deviations - Matrix Form

	(1)	(2)	(3)	(4)	(5)	(6)
(1) UA-over	1					
(2) UA-side	1	1				
(3) PA-over	1	1	1			
(4) PA-side	1	1	1	1		
(5) FA-over	1	1	1	1	1	
(6) FA-side	1	1	1	1	1	1

Figure 2. Example of a SWORD judgment matrix. This matrix corresponds to the evaluation form displayed in Figure 1.

Abbildung 10

Darstellungen der SWORD Beurteilungsmatrix, aus Vidulich et al. (1991)

2.1.16 VACP bzw. McCracken-Aldrich-Approach

	Visual, Auditory, Cognitive, Psychomotor
Autoren	McCracken und Aldrich (1984); Aldrich und Szabo (1986)
Verfahren	Fremdrating, qualitative & quantitative Methodik, objektiv
Datenerhebung	Post Trial, Aufgabenanalyse nach Beobachtung, Paper & Pencil oder mit <i>IM-PRINT</i> -Software (Mitchell, 2000), je nach Aufgabenlänge zeitaufwändig
Durchführung	<p>Nach einer durchgeführten Aufgabe oder Mission wird eine Timeline des Ablaufs erstellt und eine Aufgabenanalyse aller ausgeführten Teilaufgaben durchgeführt. Dadurch werden die einzelnen Teilaufgaben identifiziert und beschrieben. Diese werden daraufhin einem der vier Anforderungsbereiche von WL (<i>visual, auditory, cognitive, psychomotor</i>) zugeordnet. Für jede der vier Dimensionen werden anhand von sieben <i>verbal anchors</i>, also einer Skala, die von 1 bis 7 Angaben über die Höhe der WL einer Aufgabe macht, standardisierte Ratings abgegeben. Dafür wird jede beschriebene Teilaufgabe mit den verbalen Ankern auf der jeweiligen Skala verglichen und einem Wert zugeordnet. Je höher der Wert ausfällt, desto höher ist die WL der Aufgabe. Zusätzlich wird die Länge jeder Teilaufgabe in Sekundenschritten erfasst (Aldrich & Szabo, 1986; Lysaght et al., 1989). Für eine Beschreibung der jeweils sieben Ankerpunkte der vier Dimensionen siehe Abbildung 11.</p> <p><u>Berechnung der WL:</u> Es werden in jeder Dimension alle Scores je zehn-Sekunden-Intervall aufsummiert. Daraus resultieren vier Werte pro Intervall.</p> <p><u>Interpretation der Werte:</u> Die Summenscores können pro Dimension interpretiert werden. Ein Wert von 7 wird als obere Grenze der Anforderungs-Kapazität angesehen, die ein Mensch erreichen kann. Werte über 8 zeigen eine Überlastung bzw. Überbeanspruchung (<i>overload</i>) an. Außerdem kann durch die kleinschrittige Dokumentation ein zeitlicher Verlauf der Belastungen und Überbeanspruchungen nachvollzogen werden.</p> <p><u>Feedback-Schleife:</u></p> <ul style="list-style-type: none"> □ Die Aufgaben bzw. die zeitlichen Abschnitte, die zu einer Überforderung führen, können durch stärkere Automatisierung des Systems entzerrt werden und dadurch die WL des Arbeitenden reduzieren.

	<p>□ Die Ergebnisse der Aufgabenanalyse können zudem zur Vorhersage der WL bei anderen Arbeitern während der Ausführung dieser Aufgabe bzw. Mission herangezogen werden.</p> <p>Laut Lysaght et al. (1989) wurden die vier Dimensionen um die Dimension <i>kinesthetic</i> ergänzt.</p>
Vorbereitung	Bei der Nutzung zur Prädiktion von WL müssen alle Schritte, die eben beschrieben wurden, vor der Aufgabendurchführung bearbeitet werden.
Validierung	<p>Das VACP Modell wird von Mitchell (2000) zusammen mit dem NASA-TLX und der SWAT als typisches Verfahren zur Erhebung der WL genannt; es soll sensitiv, wenig intrusiv und gut auf andere Anwendungsbereiche übertragbar sein (vgl. Mitchell, 2000, S. 6). Sie nennt allerdings keine Kennwerte oder Quellen.</p> <p>Yee et al. (2007) erfassten mittels VACP Modell – das mithilfe der U.S. Army Software IMPRINT ausgewertet wurde – Anforderungen während des Fahrens bei Sekundäraufgaben (zusätzliches Telefonieren, CD wechseln, etc.)</p> <ul style="list-style-type: none"> - $N = 96$, gleichverteilt auf Männer und Frauen aus drei Altersklassen - Zuordnung der Sekundäraufgaben zu den vier VACP Dimensionen - Deskriptive Auswertung der WL aller Sekundäraufgaben - Höchste Korrelation der visuellen und kognitiven Anforderungen auf das Fahren: $r = .68$; visuell und psychomotorisch: $r = .48$
Bewertung	<p>Wird trotz mangelnder Validierung häufig verwendet, um die jeweiligen Anforderungen (<i>visual, auditory, cognitive, psychomotor</i>) in komplexen Vorhersagemodellen von WL vorherzusagen (Cain, 2007).</p> <p>➔ Interessanter Ansatz, aber ohne konkrete Ergebnisse aus Validierungsstudien schwer zu bewerten und daher nur eingeschränkt empfehlenswert</p>
Verfügbarkeit	McCracken und Aldrich (1984); Aldrich und Szabo (1986)

Workload Component Scales

SCALE VALUE	DESCRIPTORS
<u>VISUAL</u>	
1	Monitor, scan, survey
2	Detect movement, change in size, brightness
3	Trace, follow, track
4	Align, aim, orient on
5	Discriminate symbols, numbers, words
6	Discriminate based on multiple aspects
7	Read, decipher text, decode
<u>AUDITORY</u>	
1	Detect occurrence of sound, tone, etc.
2	Detect change in amplitude, pulse rate, pitch
3	Comprehend semantic content of message
4	Discriminate sounds on the basis of signal pattern, pitch, pulse rate, amplitude
<u>COGNITIVE</u>	
1	Automatic [simple association]
2	Sign/signal recognition
3	Alternative selection
4	Encoding/decoding, recall
5	Formulation of plans [projecting action sequence, etc.]
6	Evaluation [consider several aspects in reaching judgment]
7	Estimation, calculation, conversion
<u>PSYCHOMOTOR</u>	
1	Discrete actuation [button, toggle, trigger]
2	Discrete adjustive [variable dial, etc.]
3	Speech using prescribed format
4	Continuous adjustive [flight controls, sensor control, etc.]
5	Manipulative [handling objects, maps, etc.]
6	Symbolic production [writing]
7	Serial discrete manipulation [keyboard entries]

Abbildung 11

Verbale Anker der vier VACP Dimensionen, aus Aldrich und Szabo (1986, S. 634)

2.1.17 VOSO

	Verbal Online Subjective Opinion
Autoren	<i>Konnten nicht gefunden werden, in manchen Artikeln wurde auf Miller (2001) verwiesen. Unserer Einschätzung nach ist sie allerdings eher nicht die Autorin dieses Verfahrens.</i>
Verfahren	Unidimensional, subjektiv
Datenerhebung	Rating wird verbal angegeben, sobald ein Ton ertönt (z. B. alle 10 Sekunden), computerbasiert, Einzeltestung, 11-stufige Skala
Durchführung	Vergleichbar mit der Overall Workload Scale. Mental Workload soll zwischen 0 und 10 eingeschätzt werden.
Vorbereitung	
Validierung	<p>Antin und Wierwille (1984) untersuchten bei $N = 30$ Studierenden die Fähigkeit des Verfahrens, die momentane WL-Belastung zu messen (<i>instantaneous mental workload, IMWL</i>).</p> <ul style="list-style-type: none"> <input type="checkbox"/> Vorhersage des VOSO Ratings über die Aufgabenbelastung zweier Aufgaben (Mittelungsaufgabe und visuelle Suchaufgabe) mit $R^2 = .31$ und $R^2 = .48$ möglich <input type="checkbox"/> Bei Mittelung über die Teilnehmer hinweg: Vorhersage des Ratings durch Aufgabenbelastung mit $R^2 = .38$ und $R^2 = .61$ möglich <p>Validierungsstudie von Charkhandaz Yeganeh et al. (2016) mit $N = 90$ männlichen Medizinstudenten</p> <ul style="list-style-type: none"> <input type="checkbox"/> VOSO und MCH wurden zur Einschätzung der WL bei einer Gedächtnisaufgabe als valide befunden: Inhaltsvalidität und konvergente Validität von VOSO und MCH <input type="checkbox"/> Gute Retest-Reliabilität für VOSO: $r = .86$ <input type="checkbox"/> (<i>Die Studie ist in Persisch verfasst, die Informationen wurden demnach, ohne tiefergehenden Einblick in die Methodik, aus dem englischen Abstract übernommen</i>) <p>Alimohammadi et al. (2020): Untersuchung der WL beim Autofahren bei $N = 48$ männlichen Fahrern in einem Fahrsimulator</p>

	<ul style="list-style-type: none"> <input type="checkbox"/> Hohe Korrelation von hohen RTs (auf auftauchende Hindernisse) und Höhe der WL gemessen über die VOSO-Ratings: $r = .58$ <input type="checkbox"/> Hohe Korrelation zwischen Spurabweichungen und VOSO: $r = .80$ ➔ VOSO ist sensitiv für Höhe der WL im Kontext von Autofahren
Bewertung	<p>Die VOSO ist besonders sensitiv bei kurzen Intervallen von mentaler WL (Miller, 2001). Möglicherweise interferiert die Erhebung allerdings mit verbalen Aufgaben, v.a. bei solchen mit Nennung von anderen Zahlen (Antin & Wierwille, 1984). Abgesehen davon handelt es sich um ein valides, reliables (Charkhandaz Yeganeh et al., 2016) und sensitives (Alimohammadi et al., 2020) Messinstrument, das einfach angewendet und ausgewertet werden kann. Es wurde im Kontext von Fahrsimulationen erprobt (Alimohammadi et al., 2020) und ist daher auch für solche Testsituationen zu empfehlen. Wie geläufig das Verfahren im Vergleich mit anderen WL-Messinstrumenten ist, wurde hier nicht berücksichtigt.</p>
Verfügbarkeit	Miller (2001)

Level of Workload:

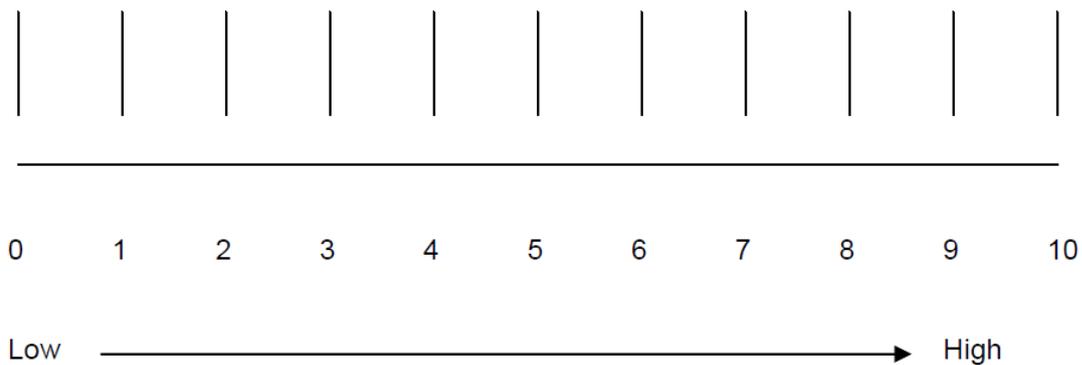


Abbildung 12
Elf-stufige Skala der VOSO, Abbildung aus Miller (2001)

2.1.18 WCI/TE

	Workload/Compensation/Interference/Technical Effectiveness
Autoren	Lysaght et al. (1989)
Verfahren	Multidimensional, subjektiv
Datenerhebung	Paper & Pencil Durchführung und computerbasierte Auswertung, Rating von 16 Zellen einer multidimensionalen Matrix
Durchführung	<p>Multidimensionale Matrix mit 16 Zellen, die geratet werden müssen und Rating der verschiedenen Aufgaben. Anschließend aufwendige computerbasierte Auswertung mit einem Ergebnis zwischen 0 und 100, das die Workload darstellt. Mehrschrittiges Verfahren (Casali & Wierwille, 1984):</p> <ul style="list-style-type: none"> • Unabhängige Bewertung von 1. WCI: <i>Workload Compensation Interference (low bis extreme)</i> und 2. TE: <i>Technical Effectiveness</i> • Kombination der Ratings (eine Zelle in der zweidimensionalen Matrix) zu einem annähernd intervallskalierten Wert • Subtraktion des Ratings von 100, um die Richtung des Wertes in ein Gesamtworkload Rating umzuwandeln. Folge: Hoher Wert = niedrige TE und hohe WCI
Vorbereitung	
Validierung	<p>Wierwille et al. (1985)</p> <ul style="list-style-type: none"> □ Sensitives Verfahren, das zwischen <i>low</i> und <i>high</i> und <i>medium</i> und <i>high</i> Load unterscheiden konnte □ Linearer Zusammenhang zwischen WCI/TE-Werten und Höhe der Load
Bewertung	Bei diesem Verfahren handelt es sich um einen sensitiven Test, der allerdings laut O'Donnell und Eggemeier (1986) nicht als direkter Test von Workload genutzt werden sollte, sondern in Ergänzung mit anderen Verfahren. Wird von Wierwille et al. (1985) als passendes Verfahren zur Erhebung von motorischer Load empfohlen (Cain, 2007).
Verfügbarkeit	Miller (2001)

	1	2	3	4	
TECHNICAL EFFECTIVENESS	Multiple Tasks Integrated				1
	Design Enhances Specific Task Accomplishment				2
	Adequate Performance Achievable; Design Sufficient to Specific Task				3
	Inadequate Performance Due to Technical Design				4
	Workload Extreme; Compensation Extreme; Interference Extreme	Workload High; Compensation High; Interference High	Workload Moderate; Compensation Moderate; Interference Moderate	Workload Low; Compensation Low; Interference Low	
	WORKLOAD/COMPENSATION/INTERFERENCE				

Abbildung 13

Multidimensionale Matrix der WCI/TE Methode, Abbildung aus Gawron (2019)

2.1.19 WP

	Workload Profile
Autoren	Tsang und Velazquez (1996)
Verfahren	Multidimensional, subjektiv
Datenerhebung	Dauer: 15-30 min., retrospektiv, Paper & Pencil, Einzeltestung, 8 Dimensionen
Durchführung	<p>8 Dimensionen:</p> <ul style="list-style-type: none"> (1) perceptual/central (2) response (3) spatial (4) verbal (5) visual (6) auditory (7) manual (8) speech <p>Für jede Dimension soll ein Rating zwischen 0 und 1 abgegeben werden, wie viele Ressourcen bzw. Aufmerksamkeit für die jeweilige Dimension verbraucht wurden. Dies erfolgt über einen Bewertungsbogen mit den verschiedenen Aufgaben in zufälliger Reihenfolge als Zeilen und den sechs WL Dimensionen als Spalten (s. Abbildung 14). Während des Ratings sind die Definitionen der Dimensionen einsehbar. In jede Zelle des Bewertungsbogen wird eine Zahl zwischen 0 (<i>keine Anforderung in dieser Dimension</i>) und 1 (<i>maximale Aufmerksamkeit</i>) eingetragen, die anzeigen soll, wieviel Aufmerksamkeitsressourcen der spezifischen Dimension in der spezifischen Aufgabe genutzt wurden. Für jede Aufgabe wird durch Summenbildung ein Gesamtrating der WL erfasst (Rubio et al., 2004; Tsang & Velazquez, 1996).</p>
Vorbereitung	Bewertungsbogen mit den jeweils genutzten Aufgaben erstellen.
Validierung	<p>Rubio et al. (2004) vergleichen die Validität von WP, TLX und SWAT.</p> <ul style="list-style-type: none"> <input type="checkbox"/> Sensitivität: Sowohl in einer single-task- als auch in einer dual-task-Bedingung signifikante Zusammenhänge mit objektiven

	<p>Veränderungen der Aufgabenschwierigkeit. Hier bessere Sensitivität als TLX und SWAT.</p> <ul style="list-style-type: none"> □ Konvergente Validität: $r = .99$ mit TLX und $r = .97$ mit SWAT □ Konkurrente Validität: $r = .73$ mit der Reaktionszeit und $r = .30$ mit einem mittleren Fehler. Allerdings geringerer Zusammenhang mit der Leistung als TLX ($r = .75$ und $r = .65$).
Bewertung	<p>Neben NASA-TLX und SWAT am häufigsten genutzt, hohe Sensitivität auch bei verschiedensten Aufgaben (Rubio et al., 2004), gute Diagnostizität und hoch reliabel (Rubio et al., 2004).</p> <p>→ Für anschließende Bewertung der Workload sehr gut geeignet, für Assessment während einer Aufgabe ungeeignet</p>
Verfügbarkeit	Rubio et al. (2004)

Workload Dimensions								
Task	Stage of processing		Code of processing		Input		Output	
	Perceptual/ Central	Response	Spatial	Verbal	Visual	Auditory	Manual	Speech
m2								
m2s1								
m2s3								
m4								
m4s1								
m4s3								
s1								
s3								

Abbildung 14
Ratingbogen des WP, Abbildung aus Rubio et al. (2004)

2.2 Planungshilfen und Empfehlungen zu WL-Erfassung

Der NASA-TLX von Sandra Hart und Kollegen ist sicherlich der am weitesten verbreitete Test zur Erfassung von WL und kann in vielerlei Hinsicht als Standard aufgefasst werden. Neue Tests müssen sich an ihm messen, und wenn nicht klar erkennbar ist, welcher Test für die gegebene Fragestellung am angemessensten ist, so spricht seine große Verbreitung für den TLX. Es gibt allerdings Entscheidungshilfen zur Auswahl eines geeigneten WL-Tests, die jetzt kurz zur Sprache kommen sollen.

Als beste Alternative zum NASA-TLX empfiehlt Cain (2007) den SWAT, als sehr geeignete Verfahren nennt er darüber hinaus VACP und DRAWS.

In Miller (2001) sind folgende Tests graphisch abgebildet:

- NASA Task Load Index (NASA-TLX)
- Subjective Workload Assessment Technique (SWAT)
- Instantaneous Self-Assessment (ISA)
- Rating Scale Mental Effort (RSME)
- Activation Scale
- Verbal Online Subjective Opinion (VOSO)
- Modified Cooper Harper Scale (MCH)
- Cooper Harper Scale
- Bedford Scale
- Analytical Hierarchy Process (AHP)
- The Workload/Compensation/Interference/Technical Effectiveness (WCI/TE)

Außerdem schlägt Sarah Miller (2001) Entscheidungsbäume vor, die dabei helfen sollen, sich für das passendste Testverfahren zu entscheiden.

Casner und Gore (2010, Figure 6) beschreiben verschiedenste Verfahren zur Erhebung von WL generell sehr anschaulich und geben hilfreiche Informationen für die Planung einer Testung, z. B.:

Technique	Prep	Time	Eyes	Ears	Hands	Absolute	Relative	Anchored
Performance								
Speed and Accuracy						x		
Activity						x		
Task Analysis	x							
Indirect								
Secondary Tasks		x	x	x	x	x		
Subjective								
<i>Numerical</i>								
ISA				x		x		
TLX		x		x		x		
Bedford			x	x		x		x
<i>Comparative</i>								
SWORD			x		x		x	
Physiological								
Heart Rate						x		
HRV						x		
Evoked Potentials						x		

Figure 6. Characteristics of workload measurement techniques considered.

Abbildung 15

Verfahren zur WL-Bestimmung laut Casner & Gore (2010)

Auch in A. J. Rehmann (1995, Table 2) ist eine informative Zusammenstellung verschiedener Verfahren zu finden:

TABLE 2. SIX COMMON SUBJECTIVE RATING SCALES JUDGED ON SEVERAL CRITERIA

Technique	Sensitivity	Reliability	Diagnosticsity	Cost/Effort Requirements	Task Time	Ease of Scoring
Analytical Hierarchy Process	High	High	Moderate	Low Cost Low Effort	Requires rating pairs of tasks	Computer scored
Bedford	High	High	Low	Low Cost Low Effort	Requires two decisions	No scoring needed
Cooper-Harper	High for psychomotor	High	Low	Low Cost Low Effort	Requires three decisions	No scoring needed
Modified Cooper-Harper	High	High	Low	Low Cost Low Effort	Requires three decisions	No scoring needed
NASA-TLX	High	High	Moderate/High	Low Cost Low Effort	Requires six ratings	Requires weighting procedure
SWAT	High	High	Moderate/High	Low Cost Low Effort	Requires prior card sort and three ratings	Requires computer scoring

(Source: Lysaght et al., 1989; ANSI/AIAA, 1992)

Abbildung 16

Verfahren zur WL-Bestimmung laut A. J. Rehmann (1995)

3 Testverfahren zur Erfassung von Situational Awareness

3.1.1 CARS

	Crew Awareness Rating Scale
Autoren	McGuinness und Foy (2000)
Verfahren	Multidimensional, subjektiv, Selbstrating
Datenerhebung	Paper & Pencil, 8 Items, 4-stufige Skala, kurz, Einzel-/Gruppentestung
Durchführung	Entwickelt für Infanterieoperationen des Militärs, Selbstrating basierend auf den drei Faktoren von SA nach Endsley (1988) und ihrer Integration. Pro Faktor zwei Aspekte (s. Abbildung 17). Ratings erfolgen auf einer vier-stufigen Skala von 1 (<i>the ideal case</i>) bis 4 (<i>the worst case</i>) (McGuinness & Ebbage, 2002). Die Datenerhebung findet nach einer Aufgabe statt.
Vorbereitung	HF Experten sollten die CARS anpassen, wenn sie in anderen Bereichen genutzt wird.
Validierung	Lee et al. (2001): NASA-TLX machte 36 % der Varianzaufklärung für CARS Werte aus (mittels multipler Regression) <ul style="list-style-type: none"> <input type="checkbox"/> fragwürdige Interrater-Reliabilität von .54, <input type="checkbox"/> aber Interrater-Übereinstimmung mit .80 als gut anzusehen <input type="checkbox"/> Angaben zu Validität sind uneindeutig bzw. nicht vorhanden
Bewertung	Vorteile: <ul style="list-style-type: none"> <input type="checkbox"/> einfache Anwendung, <input type="checkbox"/> wenig vorheriges Training nötig Nachteile: <ul style="list-style-type: none"> <input type="checkbox"/> fragliche Konstruktvalidität, <input type="checkbox"/> wenig Validierung, <input type="checkbox"/> seltene Anwendung, <input type="checkbox"/> post-trial Erhebung der SA – evtl. mit Workload konfundiert und Gedächtniseffekte! (Salmon et al., 2006)

	→ Gut anwendbares Instrument, um SA in Infanterieoperationen zu erheben, aber Probleme mit diagnostischer Güte
Verfügbarkeit	Abgedruckt in Gatsoulis (2008), Appendix B, ab S. 174

TABLE 3.3

Definitions of CARS Rating Scales

Perception – the assimilation of new information

1. The content of perception – is it reliable and accurate?

2. The processing of perception – is it easy to maintain?

Comprehension – the understanding of information in context

3. The content of comprehension – is it reliable and accurate?

4. The processing of comprehension – is it easy to maintain?

Projection – the anticipation of possible future developments

5. The content of projection – is it reliable and accurate?

6. The processing of projection – is it easy to maintain?

Integration – the synthesis of the above with one's course of action

7. The content of integration – is it reliable and accurate?

8. The processing of integration – is it easy to maintain?

Abbildung 17

Definition der CARS Skalen, Abbildung aus Gawron (2019)

3.1.2 CLSA

	China Lake Situational Awareness
Autoren	Adams (1998)
Verfahren	Unidimensional, subjektiv, Selbstrating
Datenerhebung	5-stufige Skala, Einzeltestung, kurz, meist computerbasiert, aber auch Paper & Pencil möglich
Durchführung	Unidimensionale, fünfstufige Ratingskala, basierend auf der fünfstufigen Bedford Workload Scale. Als Verfahren wurde es entwickelt, um SA während Flügen zu erfassen und um weitere WL-Verfahren und Messungen wie z. B. von Head-Up-Display-Kameras zu ergänzen. Es ist nicht als eigenständiges SA-Testverfahren angedacht. Subjektive Ratings werden von Piloten während Übungsflügen abgegeben von 1 (<i>very good</i>) bis 5 (<i>poor</i>) (Adams, 1998). Anwendung laut Adams (1998) auf Übungsflüge beschränkt.
Vorbereitung	Übungsflüge im Simulator und CLSA Ratings währenddessen umsetzen.
Validierung	<p>Olson (2003): Bei $N = 12$ Pilotenschülern wurde nach Einführung des TSAS (Tactile Situational Awareness System) eine geringere WL (über MCH und OW erfasst) und eine verbesserte SA mit CLSA valide gemessen.</p> <p>Jennings et al. (2004): Ebenfalls Verwendung von CLSA und MCH, um das TSAS zu validieren. Ähnliche Ergebnisse wie bei Olson (2003).</p> <p>G. S. Taylor et al. (2017): Bei $N = 8$ erfahrenen Militärpiloten hingen niedrigere WL Werte (über Bedford Scale gemessen) zusammen mit besserer SA, aber es erfolgte keine Angabe von Kennwerten.</p>
Bewertung	<p>Es liegt eine Validierungsstudie vor, die online allerdings nicht zu finden ist:</p> <p><i>Adams, S.R., Kane, R. and Bates R. (1998). Validation of the China Lake Situational Awareness scale with 3D SART and S-CAT. China Lake, CA: Naval Air Warfare Center Weapons Division (452330D).</i></p> <p>Gawron et al. (1999) empfehlen den Gebrauch der CLSA zur Erfassung von SA.</p>

	→ Konzeptuell schlüssige Überlegung, wird häufig eingesetzt; ohne weitere Informationen zur Validität und anderen Kennwerten ist eine angemessene Bewertung dieses Verfahrens schwierig.
Verfügbarkeit	Abgedruckt in Olson (2003), Appendix F, S. 168

SA SCALE VALUE	CONTENT
VERY GOOD 1	<ul style="list-style-type: none"> • Full knowledge of a/c energy state/tactical environment/mission; • Full ability to anticipate/accommodate trends
GOOD 2	<ul style="list-style-type: none"> • Full knowledge of a/c energy state/tactical environment/mission; • Partial ability to anticipate/accommodate trends; • No task shedding
ADEQUATE 3	<ul style="list-style-type: none"> • Full knowledge of a/c energy state/tactical environment/mission; • Saturated ability to anticipate/accommodate trends; • Some shedding of minor tasks
POOR 4	<ul style="list-style-type: none"> • Fair knowledge of a/c energy state/tactical environment/mission; • Saturated ability to anticipate/accommodate trends; • Shedding of all minor tasks as well as many not essential to flight safety/mission effectiveness
VERY POOR 5	<ul style="list-style-type: none"> • Minimal knowlege of a/c energy state/tactical environment/mission; • Oversaturated ability to anticipate/accommodate trends; • Shedding of all tasks not absolutely essential to flight safety/mission effectiveness

Abbildung 18

Definition der fünf Ratingstufen von CLSA, Quelle der Abbildung: [Researchgate](#)

3.1.3 Cranfield SAS

	Cranfield Situation Awareness Scale
Autoren	Dennehy (1997)
Verfahren	Multidimensional, subjektiv: Selbst- und Fremdratings
Datenerhebung	Paper & Pencil, geringe Dauer, 5 Skalen, Gruppen- & Einzeltestung möglich
Durchführung	<p>Die Cranfield-SAS kann als Fremd- oder Selbsteinschätzung der SA von PilotenschülerInnen während oder nach einer Flugsimulation eingesetzt werden.</p> <p>Multi-dimensionales Rating Instrument, das die SA über folgende Skalen erhebt:</p> <ol style="list-style-type: none"> (1) pilot knowledge (2) understanding and anticipation of future events (3) management of stress, effort and commitment (4) ability to attend, perceive, assimilate and access information (5) overall awareness <p>Für die Inhalte der einzelnen Skalen s. Dennehy (1997), S.2.</p> <p>Ratings je Frage werden von (1) <i>unacceptable</i> bis (9) <i>excellent</i> abgegeben.</p> <p>Anschließend werden alle Ratings pro Skala aufsummiert. Je höher der Summenscore ausfällt, desto höher ist die SA.</p> <p>Kurzform und Langform vorhanden (Gawron, 2019).</p> <p>Anwendbar in diesen sechs Trainingsabschnitten: general aircraft handling, navigation, instrument flying (basic), instrument flying (airways), night flying, commercial pilot flight test.</p>
Vorbereitung	Mit Inhalten der Skalen, den Fragen und dem Bewertungssystem vertraut machen.
Validierung	Die Berechnung und das Zutreffen der Gütekriterien wird beschrieben, es fehlen allerdings Werte (Dennehy, 1997).
Bewertung	Geringe Intrusivität, zur Evaluierung realer Systeme (keine Simulationen) geeignet (Jeannot et al., 2003).

	<p>Einfach anwendbares Instrument zur Fremd- & Selbsteinschätzung, aber keine Validierungsstudien (Salmon et al., 2006).</p> <p>➔ Umfangreiches Verfahren durch die Kombination von Selbst- und Fremdratings, scheinbar valide und reliable, ohne konkrete Kennwerte ist die diagnostische Güte allerdings schwer zu bewerten.</p>
Verfügbarkeit	<p>Kurz- und Langform der Cranfield SAS sind beispielhaft in Dennehy (1997) abgedruckt.</p>

3.1.4 MARS

	Mission Awareness Rating Scale
Autoren	Matthews et al. (2002)
Verfahren	Multidimensional, subjektiv, Selbstrating
Datenerhebung	Paper & Pencil, je 4 Items zu SA & WL, 4-stufige Skala, Einzel-/Gruppentestung
Durchführung	Testverfahren auf Grundlage der CARS. Die MARS wurde speziell für Messungen von SA in Infanterieübungen in realen Umgebungsbedingungen, und nicht in Simulationen, entwickelt. Besteht aus zwei Sets an Fragen: Vier Fragen zu SA, vier zu WL.
Vorbereitung	Formulierung der Rating Scales je nach Art der Mission
Validierung	Es sind wenig Validierungsstudien zu finden. Laut Matthews et al. (2011) verfügt die MARS über prädiktive Validität, untersucht bei norwegischen Kadetten.
Bewertung	Gleiche Vor- & Nachteile wie CARS.
Verfügbarkeit	Ausführlich in Matthews et al. (2002) für das o.g. Beispiel abgebildet.

Content subscales

Please rate your ability to identify mission-critical cues in this mission.

- Very easy- able to identify all cues
- Fairly easy – could identify most cues
- Somewhat difficult – many cues hard to identify
- Very difficult – had substantial problems identifying most cues

How well did you understand what was going on during the mission?

- Very well – fully understood the situation as it unfolded
- Fairly well – understood most aspects of the situation
- Somewhat poorly – had difficulty understanding much of the situation
- Very poorly – the situation did not make sense to me

How well could you predict what was about to occur next in the mission?

- Very well – could predict with accuracy what was about to occur
- Fairly well – could make accurate predictions most of the time
- Somewhat poor – misunderstood the situation much of the time
- Very poor – unable to predict what was about to occur

How aware were you of how to best achieve your goals during this mission?

- Very aware – knew how to achieve goals at all times
- Fairly aware – knew most of the time how to achieve mission goals
- Somewhat unaware – was not aware of how to achieve some goals
- Very unaware – generally unaware of how to achieve goals

Abbildung 19

Items zur Erhebung von SA durch die MARS, Abbildung aus Cummings et al. (2008)

3.1.5 QASA

	Quantitative Analysis of Situation Awareness
Autoren	Edgar et al. (2000)
Verfahren	Multidimensional, Subjektiv & Performance measure
Datenerhebung	Paper & Pencil, kurz, mind. 20 Items, zwei Dimensionen, Einzeltestung
Durchführung	<p>Hypothese: eine gute SA liegt vor, wenn für die Situation relevante richtige von falschen Informationen unterschieden werden können.</p> <p>Die Anwendung von QUASA resultiert in zwei verschiedenen Scores: SA und Information Bias (IB).</p> <p>Ablauf: Richtige und plausible falsche Aussagen werden präsentiert, Probanden müssen entscheiden, ob die Aussage richtig oder falsch ist. Die subjektive Sicherheit (IB) des Urteils wird ebenfalls erfasst. Mittels Signalentdeckungstheorie wird analysiert, wie gut Probanden in diesen Unterscheidungen sind. SA wird als Anteil (von 0 bis 100) der korrekten Antworten berechnet (Edgar et al., 2000). IB kann von -100 ("Individual accepts all information as true and is therefore basing decisions on much false information") bis +100 ("Individual rejects all information – including true info") (Edgar et al., 2000, S. 29).</p>
Vorbereitung	<p>Szenario überlegen und anschließend eine Situationsanalyse von bzw. mit Experten durchführen lassen, damit die Ergebnisse repräsentativ für die jeweils untersuchte Situation sind.</p> <p>Mindestens 20 Wahr/Falsch-Aussagen erstellen, von denen etwa 50% in der Situation zutreffen.</p>
Validierung	Untersuchung der Gütekriterien von QUASA anhand einer Studie mit Feuerwehrleuten (Edgar et al., 2018), in der 42 Items von einem Experten erstellt wurden. Interrater-Reliabilität der Items mit .85 als sehr gut anzusehen.
Bewertung	<p>Vorteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Objektive Werte und subjektive Einschätzung von SA (Salmon et al., 2006)

	<ul style="list-style-type: none"> <input type="checkbox"/> Gibt durch Auswertung über Signalentdeckungstheorie zusätzlich Aufschluss über Ausprägung, sprich die Richtung von Urteilsverzerrungen (McGuinness, 2004). <input type="checkbox"/> Auf viele Kontexte anwendbar, wenn Situationsanalyse und Itemerstellung von Experten realisierbar sind <p>Nachteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Validierungsstudien zwar vorhanden, aber keine Werte der Gütekriterien angegeben (Edgar et al., 2000) <input type="checkbox"/> Erfassung der SA intrusiv zur Hauptaufgabe (Salmon et al., 2006)
Verfügbarkeit	Beispielhaft abgedruckt in Edgar et al. (2000).

3.1.6 SAGAT

	Situation Awareness Global Assessment Technique
Autorin	Endsley (1988), Endsley (1990)
Verfahren	Multidimensional bzw. Leistungsmessung, objektiv
Datenerhebung	Während einer Simulation, computerbasiert, Dauer eines Messzeitpunkts: zwischen 30-120 Sek., Einzeltestung
Durchführung	In unregelmäßigen Abständen wird ein simuliertes, gerade operierendes System angehalten (z.B. Flugsimulator). Der Operator wird nach relevanten Informationen über die momentane Situation des Systems befragt. Dabei werden alle drei Level (<i>perception, comprehension, projection</i>) von SA nach Endsley (1988) abgedeckt. Die Antworten werden anschließend mit den Daten der Simulator-Software verglichen. Dieser Vergleich zwischen der tatsächlichen und der wahrgenommenen Situation gibt Auskunft über die Höhe der SA des Operators: je kongruenter die Daten, desto höher die SA (Endsley, 1990).
Vorbereitung	Der Nutzung der SAGAT geht eine ausführliche Aufgabenanalyse voraus. Für folgende Felder bestehen bereits Befragungen: fighter aircraft, bomber aircraft, commercial aircraft, air traffic control, maintenance systems, nuclear power. Für alle anderen Bereiche muss die Befragung angepasst werden (Endsley et al., 1998).
Validierung	Erhebung von SAGAT nicht nachteilig für Leistung von Piloten ($N = 6$) (Snow & Reising, 2000). Das Verfahren besitzt eine gute Diagnostizität und Augenscheinvalidität (Snow & Reising, 2000). Loft et al. (2015): Bei $N = 171$ Studierenden wurde für die SAGAT prädiktive Validität gefunden, höher als für SPAM. Aber die Erhebung von SAGAT führte zu leicht erhöhter WL (gemessen über NASA-TLX und ATWIT). Zahlreiche weitere Validierungsstudien, siehe Abb. 20
Bewertung	Vorteile: <ul style="list-style-type: none"> <input type="checkbox"/> Objektive Daten <input type="checkbox"/> Hohe Inhaltsvalidität

	<ul style="list-style-type: none"> □ Prädiktive Validität z.B. Vorhersage von Performanz in Flugsimulation eines Kampfes (Endsley, 1990; Endsley et al., 1998) und auch über eine Vielzahl verschiedener Studien hinweg gute Prädiktion der Performanz (Endsley, 2019) □ Gute Sensitivität und Diagnostizität, hohe Reliabilität (Salmon et al., 2006), bessere Sensitivität als SPAM (Endsley, 2019) □ Erfasst SA sowohl bei hohem als auch bei niedrigem WL (Endsley, 2019) □ Keine Intrusivität von freeze probes, keine negativen Auswirkungen auf die Performanz (Endsley, 2019) □ Keine übermäßige Verzerrungen durch Gedächtniseffekte trotz freeze probes (Endsley, 2019) □ Wird häufig kombiniert mit physiologischen Messungen von SA (z.B. Eye-Tracking), die während der Bearbeitung der Aufgabe nicht-intrusiv die SA erfassen können (Zhang, 2023). Die meisten hier aufgeführten Testverfahren sind mit physiologischen Maßen kombinierbar. Für weitere Informationen zu physiologischen Messmethoden siehe Boumann, Cyrol, Biella und Sammito (2022). <p>Nachteile:</p> <ul style="list-style-type: none"> □ Anspruchsvolle Vorbereitung und Durchführung (Snow & Reising, 2000) ➔ Häufig untersuchtes und validiertes Testverfahren mit guten Kennwerten; eines der Standard-Testverfahren für SA (Salmon et al., 2006)
Verfügbarkeit	Ein ganzes Kapitel dazu in Endsley und Garland (2008).

Table 6

Reliability and Validity Terms, Definitions Correlated to SAGAT and Relevant Citations

Term	Definition	Demonstrated in SAGAT	Evidence (Citation) Recent Healthcare Literature
Face Validity	Instrument appears to measure what it is supposed to	Items derived by GDTA	Endsley, 2000; Gardner et al., 2017; Lavoie et al., 2016; Morgan et al., 2015; Salmon et al., 2009
Content Validity	Items that make up the instrument adequately represent the variable being measured.	Items derived by GDTA	Endsley, 2000; Gardner et al., 2017; Lavoie et al., 2016; Morgan et al., 2015; Salmon et al., 2009
Criterion Validity	Degree of correlation between scores on an instrument and an established standard <u>Predictive validity</u> Tool predicts observed measure in the future	GDTA based on established standards as agreed upon by content experts.	Crozier et al., 2014; Endsley, 2000; Gardner et al., 2017; Hogan et al., 2016; Lavoie et al., 2016
	<u>Concurrent validity</u> : correlation between scores on instrument with external criterion measured at same time		
Construct Validity	Instrument measures what it purports to measure	Established by the content experts	Crozier et al., 2014; Endsley, 2000; Gardner et al., 2017; Hogan et al., 2016
Sensitivity & Specificity	Sensitivity = correctly finds true positive cases Specificity = correctly identifies untrue cases		Endsley, 2000; Hogan et al., 2006; Lavoie et al., 2016
Test-Retest Reliability	Results repeated under separate occasions, and/or different test giver, subjects, scenario		Endsley, 2000; Gardner et al., 2017

Note: Definitions of terms from Polit and Beck (2012)

Abbildung 20

Übersicht von Validierungsstudien zu SAGAT, Abbildung aus Dishman (2019)

3.1.7 SALIANT

	Situational Awareness-Linked Instances Adapted to Novel Tasks
Autoren	Bowers et al. (1998); Muñiz et al. (1998)
Verfahren	Multidimensional, objektiv, Expertenrating
Datenerhebung	<i>Performance Measure</i> bzw. -Bewertung, 24 Items, 5 Dimensionen, Rating nach trifft zu/ trifft nicht zu, Einzeltestung von Teammitgliedern, computerbasiert
Durchführung	<p>Messung von SA in Teams, ursprüngliche Anwendung in der Luftfahrt, aber gut auf andere Bereiche übertragbar.</p> <p>Annahme: Die Kombination verschiedener Verhaltensweisen in einem speziellen Teamkontext spiegelt größtenteils die Team-SA wider.</p> <p>Auf Grundlage empirischer Erkenntnisse zu SA-relevanten Verhaltensweisen im Team entwickelten die Autoren eine Checklist, die Verhalten beschreibt, das in fünf verschiedene Kategorie fällt (s. auch Abb. 21):</p> <ol style="list-style-type: none"> (1) Demonstrated Awareness of Surrounding Environment (2) Recognized Problems (3) Anticipated a Need for Action (4) Demonstrated Knowledge of Tasks (5) Demonstrated Awareness of Information <p>Das gezeigte Verhalten der Teammitglieder wird per Checklist danach geratet, ob die relevanten Verhaltensweisen gezeigt wurden oder nicht und anschließend als Verhaltensmuster pro Team ausgewertet.</p>
Vorbereitung	<ol style="list-style-type: none"> 1. Relevante Verhaltensindikatoren für Team-SA recherchieren 2. Entwicklung eines Testszenarios 3. Konkrete Verhaltensindikatoren pro Simulation festlegen 4. Skript mit standardisierten Instruktionen verfassen 5. Checkliste für 3. anlegen
Validierung	<p>Bowers et al. (1998): $N = 30$ Piloten absolvierten eine Flugsimulation und wurden anschließend geratet.</p> <ul style="list-style-type: none"> - Moderate bis sehr gute Interrater-Reliabilität: $r = .70$ bis $.96$

	- Prädiktive Validität insb. für Problemlösefähigkeiten gefunden
Bewertung	<p>Vorteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Verhaltensaspekte sind gut und nicht-intrusiv messbar. <input type="checkbox"/> Gute Interrater-Reliabilität (Bowers et al., 1998; Grosjean & Terrier, 1999) <input type="checkbox"/> SALIANT erzielte höhere Varianzaufklärung im Kontext von Team-SA als SAGAT (Bowers et al., 1998). <input type="checkbox"/> Eines der wenigen Verfahren zur Messung der Team-SA (Jeannot et al., 2003). <p>Nachteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Kognitive Prozesse werden vernachlässigt. <input type="checkbox"/> Es werden nur Verhaltensweisen bewertet, die für hohe SA sprechen, solche für geringe SA fehlen. <input type="checkbox"/> Aufwendige Entwicklung einer simulativen Operation oder Aufgabe durch Experten, damit Team-SA valide erfasst werden kann (Jeannot et al., 2003).
Verfügbarkeit	Ausführliche Beschreibung des Vorgehens und Beispielformulare in Muñiz et al. (1998)

Demonstrated Awareness of Surrounding Environment
<ul style="list-style-type: none"> • Monitored environment for changes, trends, abnormal conditions (Ref 20) • Demonstrated awareness of where he/she was (Ref 16)
Recognized Problems
<ul style="list-style-type: none"> • Reported problems (Ref 19; Ref 18) • Located potential sources of problem (Ref 19) • Demonstrated knowledge of problem consequences (Ref 20) • Resolved discrepancies (Ref 14) • Noted deviations (Ref 19)
Anticipated a Need for Action
<ul style="list-style-type: none"> • Recognized a need for action (Ref 20; Ref 19; Ref 18) • Anticipated consequences of actions and decisions (Ref 20) • Informed others of actions taken (Ref 21) • Monitored actions (self & others) (Ref 19)
Demonstrated Knowledge of Tasks
<ul style="list-style-type: none"> • Demonstrated knowledge of tasks (Ref 14) • Exhibited skilled time sharing attention among tasks (Ref 14) • Monitored workload (self & others) (Ref 20) • Shared workload within station (Ref 16) • Answered questions promptly (Ref 20)
Demonstrated Awareness of Information
<ul style="list-style-type: none"> • Communicated important information (Ref 16) • Confirmed information when possible (Ref 16; Ref 20) • Challenged information when doubtful (Ref 19; Ref 21; Ref 16) • Re-checked old information (Ref 17) • Provided information in advance (Ref 19; Ref 14) • Obtained information of what is happening (Ref 18) • Demonstrated understanding of complex relationships (Ref 16; Ref 14) • Briefed status frequently (Ref 20; Ref 14)

Abbildung 21

Dimensionen und Items zur Erfassung der Team-SA mittels SALIANT, Abbildung aus Muñiz et al. (1998)

3.1.8 SALSA

	Situation Awareness bei Lotsen der Streckenflugkontrolle im Kontext von Automatisierung
Autoren	Hauss und Eyferth (2003)
Verfahren	Multidimensional, objektiv, freeze probe bzw. performance measure
Datenerhebung	Simulator-/computerbasiert, mittellange Dauer, abhängig von Primäraufgabe, 15 Items
Durchführung	<p>Freeze probe Technik, findet Anwendung in ATC. 15 verschiedene Aspekte werden abgefragt, z.B. <i>flight level, ground speed, heading, vertical tendency, conflict and type of conflict</i>.</p> <p>Die Flugsimulation wird unterbrochen, indem der Radar-Bildschirm eingefroren wird, ein Flugzeug wird hervorgehoben und es werden alle 15 Aspekte zu diesem Objekt abgefragt. Ein Expertenrating bestimmt im Nachhinein, wie relevant jeder der 15 Aspekte in der jeweiligen Testsituation war. Nur die relevanten Items werden bei der Auswertung berücksichtigt.</p>
Vorbereitung	Flugsimulator und -simulation nötig
Validierung	Valides und reliables Verfahren (Hauss & Eyferth, 2003)
Bewertung	<p>Vorteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Daten werden während einer Simulation erhoben, also keine Gedächtniseffekte <input type="checkbox"/> Geeignetes Verfahren zur Erfassung von SA in ATC-Kontext (Hauss & Eyferth, 2003) <p>Nachteile laut Salmon et al. (2006):</p> <ul style="list-style-type: none"> <input type="checkbox"/> Simulator nötig <input type="checkbox"/> Kann Ausführung der Primäraufgabe stören <input type="checkbox"/> Seltener Gebrauch <input type="checkbox"/> Wenig validiert <p>➔ Gegenüber anderen Verfahren eher weniger nützlich</p>
Verfügbarkeit	Übersicht der 15 Fragen in Hauss und Eyferth (2003).

3.1.9 SARS

	Situational Awareness Rating Scale
Autoren	Wayne et al. (1994)
Verfahren	Multidimensional, subjektiv und objektiv; 3 Formen: Selbstrating, Rating durch Supervisor (s. SASRF) und durch Peers
Datenerhebung	Paper & Pencil, 31 Items, 8 Kategorien, 6-stufige Skala, kurze Durchführung, Einzel-/Gruppentestung
Durchführung	<p>Entwickelt für die militärische Luftfahrt. Subjektives Rating der Performanz auf einer sechs-stufigen Skala (von <i>acceptable</i> bis <i>outstanding</i>) über 31 Items, die in acht Kategorien fallen:</p> <ol style="list-style-type: none"> (1) general traits (2) tactical game plan (3) system operation (4) communication (5) information interpretation (6) tactical employment – behavioral (7) tactical employment – visual (8) tactical employment – general <p>Für eine Übersicht der 31 Items siehe Abb. 22. Die acht Kategorien und insgesamt 31 Verhaltensweisen wurden auf Grundlage von Interviews mit erfahrenen F-15 Piloten entwickelt (Salmon et al., 2006; Waag & Houck, 1996).</p>
Vorbereitung	Simulation muss vorbereitet werden, auf die sich die Ratings anschließend beziehen.
Validierung	<p>Der beste Faktor zur Vorhersage von SA bei Pilotinnen ist Flugerfahrung (Carretta et al., 1996).</p> <p>Wayne et al. (1994): Untersuchung von $N = 205$ erfahrenen USAF F-15C Piloten</p> <ul style="list-style-type: none"> <input type="checkbox"/> Erhoben wurden Selbst-, Peer-Ratings und Ratings durch Supervisoren (die aus den erfahrensten Piloten bestanden) <input type="checkbox"/> Hohe interne Konsistenz: Cronbach's Alpha = .95 bis .99 <input type="checkbox"/> Hohe Interrater-Reliabilität: .88 bis .97 <p>Waag und Houck (1996): Untersuchten $N = 239$ USAF F-15 Piloten</p>

	<ul style="list-style-type: none"> <input type="checkbox"/> hohe interne Konsistenz: Cronbach's Alpha = .97 bis .99 <input type="checkbox"/> hohe Interrater-Reliabilität: $r = .84$ <input type="checkbox"/> hohe Korrelationen zwischen Peer- und Supervisor-Ratings ($r = .85$ bis $.87$) <input type="checkbox"/> hohe Korrelationen von Peer- und Supervisor-Ratings mit Selbstrating ($r = .50$ bis $.58$)
Bewertung	<p>Vorteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> sehr ökonomisch, <input type="checkbox"/> hohe interne Konsistenz, <input type="checkbox"/> schnell anwendbar ohne viel Trainingsaufwand, <input type="checkbox"/> nicht intrusiv zur Aufgabe <p>Nachteil:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Daten werden erst nach den Trials erhoben (Salmon et al., 2006) <p>➔ Speziell für SA bei Militärpiloten entwickelt, praktisch in diesem Bereich und umfangreiche Beurteilung durch die drei verschiedenen Ratingformen</p>
Verfügbarkeit	In Wayne et al. (1994)

TABLE I. ITEMS AND CATEGORIES USED IN SARS.

<p>1. GENERAL TRAITS Discipline Decisiveness Tactical knowledge Time-sharing ability Spatial ability Reasoning ability Flight management</p>	<p>5. INFORMATION INTERPRETATION Interpreting vertical situation display Interpreting threat warning system Ability to use controller information Integrating overall information Radar sorting Analyzing engagement geometry Threat prioritization</p>
<p>2. TACTICAL GAME PLAN Developing plan Executing plan Adjusting plan on-the-fly</p>	<p>6. TACTICAL EMPLOYMENT-BVR Targeting decisions Fire-point selection</p>
<p>3. SYSTEM OPERATION Radar Tactical electronic warfare system Overall weapons system proficiency</p>	<p>7. TACTICAL EMPLOYMENT-VISUAL Maintain track of bogeys/friendlies Threat evaluation Weapons employment</p>
<p>4. COMMUNICATION Quality (brevity, accuracy, timeliness) Ability to effectively use information</p>	<p>8. TACTICAL EMPLOYMENT-GENERAL Assessing offensiveness/defensiveness Lookout Defensive reaction Mutual support</p>

Abbildung 22

Items und Kategorien der SARS, Abbildung aus Wayne et al. (1994) (S. A14)

3.1.10 SART und 3-D SART

	Situation Awareness Rating Technique
Autoren	R. M. Taylor (1990)
Verfahren	Unidimensional, subjektiv, Selbstrating-Fragebogen
Datenerhebung	Paper & Pencil, 10 Items, 7-stufige Skala, kurze Durchführungsdauer, Einzel- und Gruppentestung
Durchführung	<p>Subjektive Messung durch Selbstbeurteilung der Operatoren auf einer unidimensionalen sieben-stufigen Skalen. Entwickelt für die (militärische) Luftfahrt.</p> <p>Annahme: Operatoren nutzen ihr Verständnis und Wissen über die Situation, um Entscheidungen zu treffen. Dieses Wissen kann bewusst gemacht und quantifiziert werden (R. M. Taylor, 1990).</p> <p>Mittels zehn Items (s. Abb. 23), die in drei Dimensionen eingeteilt werden, wird die SA erhoben.</p> <p><u>3-D SART</u>: Als Kurzform können auch die zehn Items als folgende drei Dimensionen zusammengefasst von <i>low</i> bis <i>high</i> geratet werden, dann allerdings in Millimeterschritten von 0 bis 100:</p> <ol style="list-style-type: none"> (1) Demand on Attentional Resources (2) Supply of Attentional Resources und (3) Understanding of the Situation (Endsley & Garland, 2008)
Vorbereitung	Selbstrating: Verfahren kann über Microsoft Excel automatisiert angewandt werden, geringer Trainingsaufwand
Validierung	<p>Endsley et al. (2000): Vergleich der Verfahren SAGAT und SART im ATC-Kontext anhand von $N = 10$ erfahrenen Fluglotsen</p> <ul style="list-style-type: none"> <input type="checkbox"/> SAGAT im Vergleich valider, reliabler und sensitiver, aber s. Bewertung
Bewertung	<p>Vorteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> SART positiv korreliert mit subjektiv gemessenem confidence level und performance measure (Endsley et al., 1998).

	<ul style="list-style-type: none"> <input type="checkbox"/> Gute Diagnostizität, korreliert mit Leistungsmessungen (Gawron, 2019). <input type="checkbox"/> Gute Sensitivität (Endsley & Garland, 2008) <p>Nachteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Messung von SA eventuell mit Messung von Workload konfundiert. <input type="checkbox"/> Keine signifikante Retest-Reliabilität (Endsley & Garland, 2008). <p>➔ Eines der bekanntesten Verfahren mit vielen Validierungsstudien, guter Diagnostizität und Sensitivität, aber fraglicher Reliabilität, das ökonomisch ist und in vielen Bereichen anwendbar (Salmon et al., 2006)</p>
Verfügbarkeit	In Endsley und Garland (2008) (sehr ausführlich); Gawron (2019)

		LOW							HIGH						
		1	2	3	4	5	6	7	1	2	3	4	5	6	7
Demand	Instability of Situation														
	Variability of Situation														
	Complexity of Situation														
Supply	Arousal														
	Spare Mental Capacity														
	Concentration														
	Division of Attention														
Under	Information Quantity														
	Information Quality														
	Familiarity														

Abbildung 23

Übersicht der zehn Items und Skala von SART, Abbildung aus Gawron (2019)

3.1.11 SASRF

	Situational Awareness Supervisory Rating Form
Autoren	Carretta et al. (1996)
Verfahren	Multidimensional, (im Idealfall) objektiv
Datenerhebung	Fremdrating, Fragebogen, 31 Items, Gruppen- und Einzeltestung möglich, i. d. R. insbesondere bei mehreren Piloten eingesetzt
Durchführung	Es handelt sich um das gleiche Verfahren wie bei der SARS, allerdings werden die Ratings hier von Supervisoren, nämlich z. B. erfahrenen F-15 Piloten, übernommen. Für alle 31 Items s. Abb. 24.
Vorbereitung	Experten bzw. Supervisoren müssen mit dem Rating vertraut sein.
Validierung (s. SARS)	<p>Der beste Faktor zur Vorhersage von SA bei Piloten ist Flugerfahrung (Carretta et al., 1996).</p> <p>Wayne et al. (1994): Untersuchung von $N = 205$ erfahrenen USAF F-15C Piloten</p> <ul style="list-style-type: none"> <input type="checkbox"/> Erhoben wurden Selbst- und Peer-Ratings sowie Ratings durch Supervisors (die aus den erfahrensten Piloten bestanden) <input type="checkbox"/> Hohe interne Konsistenz: Cronbach's Alpha = .95 bis .99 <input type="checkbox"/> Hohe Interrater Reliabilität: $r = .88$ bis .97 <p>Waag und Houck (1996): Untersuchten 239 USAF F-15 Piloten</p> <ul style="list-style-type: none"> <input type="checkbox"/> hohe interne Konsistenz: Cronbach's Alpha = .97 bis .99 <input type="checkbox"/> hohe Interrater-Reliabilität: $r = .84$ <input type="checkbox"/> hohe Korrelationen zwischen Peer- und Supervisor-Ratings ($r = .85$ bis .87) <input type="checkbox"/> hohe Korrelationen von Peer- und Supervisor-Ratings mit Selbstrating ($r = .50$ bis .58)
Bewertung	→ vielversprechende Ergebnisse, spezifische Anwendungsvoraussetzungen (Simulator, Experten als Rater), nützlich im Kontext von ATC, insbesondere bei der Beurteilung der SA von Piloten geeignet
Verfügbarkeit	In Carretta et al. (1996) und Gawron (2019)

Rater ID#: _____ Pilot ID#: _____						
Item Ratings	Relative Ability Compared with Other F-15C Pilots					
	Acceptable		Good		Outstanding	
	1	2	3	4	5	6
General Traits 1. Discipline 2. Decisiveness 3. Tactical knowledge 4. Time-sharing ability 5. Reasoning ability 6. Spatial ability 7. Flight management						
Tactical Game Plan 8. Developing plan 9. Executing plan 10. Adjusting plan on the fly						
System Operation 11. Radar 12. TEWS 13. Overall weapons system proficiency						
Communication 14. Quality (brevity, accuracy, timeliness, completeness) 15. Ability to effectively use comm/information						
Information Interpretation 16. Interpreting VSD 17. Interpreting RWR 18. Ability to effectively use AWACS/GCI 19. Integrating overall information (cockpit displays, wingman comm, controller comm) 20. Radar sorting 21. Analyzing engagement geometry 22. Threat prioritization						
Tactical Employment-BVR Weapons 23. Targeting decision 24. Fire-point selection						
Tactical Employment-Visual Maneuvering 25. Maintain track of bogeys/friendlies 26. Threat evaluation 27. Weapons employment						
Tactical Employment-General 28. Assessing offensiveness/defensiveness 29. Lookout (VSD interpretation, RWR monitoring, visual lookout) 30. Defensive reaction (chaff, flares, maneuvering, etc.) 31. Mutual support						
Overall situational awareness ^a						
Overall fighter ability						

^a Items 1 through 31 are used for supervisory ratings. The overall fighter ability and situational awareness items are completed by both supervisors and peers. (Carretta et al., 1996, pp. 40-41).

Abbildung 24

Übersicht aller 31 Items der SASRF, aus Gawron (2019), S. 172-173

3.1.12 SA-SWORD

	Situational Awareness Subjective Workload Dominance
Autoren	Vidulich und Hughes (1991)
Verfahren	Subjektiv, <i>Paired Comparison</i> Verfahren
Datenerhebung	Paper & Pencil, geringe Durchführungsdauer, Einzel- & Gruppentestung
Durchführung	Adaption des Workload Testverfahrens „SWORD“ (Vidulich et al., 1991) für SA-Messungen. <i>Paired-comparison</i> Verfahren: Entscheidungsmatrix mit allen möglichen Mehrfachvergleichen wird für die relevanten Bereiche aufgestellt und für jeden Vergleich wird ein Urteil abgegeben, wie hoch die SA in einem Bereich im Vergleich zu einem anderen Bereich aussieht (Snow & Reising, 2000; Vidulich & Hughes, 1991).
Vorbereitung	Bedeutung von SA genau definieren, damit teilnehmende Personen diese einheitlich bewerten.
Validierung	Validierungsstudie von Vidulich und Hughes (1991) mit $N = 10$ erfahrenen Piloten: <ul style="list-style-type: none"> <input type="checkbox"/> Gute Interrater Reliabilität: $r = .71$ <input type="checkbox"/> Sensitives und nützliches Verfahren, mit guter Augenscheinvalidität und Akzeptanz durch Nutzer, allerdings nur anhand einer kleinen Stichprobe untersucht! <p>Snow und Reising (2000): Vergleich zwischen SAGAT und SA-SWORD (bei $N = 12$ Piloten) ergab eine höhere Sensitivität für SA-SWORD.</p>
Bewertung	Vorteile: <ul style="list-style-type: none"> <input type="checkbox"/> einfach in der Durchführung <input type="checkbox"/> ökonomisch <input type="checkbox"/> kann in vielen Bereichen genutzt werden <input type="checkbox"/> sehr nützlich beim Vergleich von zwei verschiedenen Konzepten <input type="checkbox"/> sensitiv <p>Nachteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Erhebung erfolgt erst post-trial (Gedächtniseffekte) <input type="checkbox"/> geringe Nutzung und wenig Validierungsstudien

	<ul style="list-style-type: none">□ kein Messverfahren zur Höhe von SA, sondern ein Vergleich, in welchem Fall die SA höher war (Salmon et al., 2006) ➔ Interessantes Konzept, je nach Anforderungen und Design eventuell sinnvoll (Vergleiche), für eine Messung der Höhe von SA aber ungeeignet
Verfügbarkeit	

3.1.13 SAVANT

	Situation Awareness Verification Analysis Tool
Autoren	Willems (2000); Willems und Heiney (2002)
Verfahren	Objektiv, <i>Performance measure</i>
Datenerhebung	Simulatorbasiert, verschiedene Zeitfenster können angewendet werden, bspw. alle 12 Sek.
Durchführung	<p>Willems (2000) und Willems und Heiney (2002) entwickelten das SAVANT und variierten und erhoben während der Testung mit Fluglotsen mehrere Variablen, u.a. Task Load. Beim SAVANT handelt es sich um eine Kombination aus den beiden Testverfahren SAGAT und SPAM (Jeannot et al., 2003).</p> <p>Es findet Anwendung in der Flugsicherung. Fluglotsinnen wird kurz (z. B. 3 Sek.) eine Frage eingeblendet, welche die Lokation zweier Flugzeuge und den jeweiligen Sektor betrifft. Es werden sowohl Radar (R-Side) als auch Daten (D-Side) von Fluglotsen genutzt (Gawron, 2019). Fragen werden auf einem extra Bildschirm gestellt. Erhoben werden die Antworten auf die jeweiligen Fragen, deren Korrektheit und die RTs (Jeannot et al., 2003).</p> <p>Ergebnis des Experiments: Kürzere RTs von Fluglotsen auf der R-Side als auf der D-Side, generell kürzere RTs bei geringerer Task Load.</p>
Vorbereitung	Umsetzung wird ausführlich in Appendix E in Willems und Heiney (2002) beschrieben.
Validierung	Langan-Fox et al. (2009): Nützlichkeit und Validität des SAVANT fraglich
Bewertung	<p>Vorteile (Langan-Fox et al., 2009):</p> <ul style="list-style-type: none"> <input type="checkbox"/> Erfragung relevanter Infos zur Messung von SA <input type="checkbox"/> Daten werden während der Situation erhoben. <p>Nachteile (Langan-Fox et al., 2009):</p> <ul style="list-style-type: none"> <input type="checkbox"/> Fragen werden auf einem separaten Bildschirm gestellt, was evtl. zu Ablenkungen führen kann <input type="checkbox"/> hoher Vorbereitungsaufwand <input type="checkbox"/> Nachweise über geringe Validität

	→ Nützliches Verfahren als Alternative zu SAGAT und SPAM in Laborstudien, hat allerdings methodische Schwächen.
Verfügbarkeit	s. Willems und Heiney (2002)

3.1.14 SPAM

	Situation Present Assessment Method
Autoren	Durso et al. (1998)
Verfahren	Objektiv, Performance measure
Datenerhebung	Simulatorbasiert, Einzeltest, Reaktionszeitmessung
Durchführung	<p>Annahme: Eine gute SA hängt nicht davon ab, alle notwendigen und situationsrelevanten Informationen immer im Gedächtnis zu haben, sondern zu wissen, wo bzw. wie diese Informationen am schnellsten abgerufen werden können (Jeannot et al., 2003).</p> <p>Auf dieser Grundlage wurde ein Verfahren entwickelt, das die Antwortlatenz auf Fragen zu Vorhersagen als repräsentativen Wert für SA misst, während alle Informationen zur Beantwortung verfügbar sind. Dadurch werden Gedächtniseffekte eliminiert (Gawron, 2019). Richtigkeit der Antwort spiegelt die SA wider und Länge der RT spiegelt die Höhe der Workload wider (Durso et al., 1998).</p> <p>Während einer Flugsimulation wird ein ATC Controller kontaktiert, die Simulation wird nicht gestoppt. Es wird eine Frage gestellt, die er beantworten soll, z.B.: „Welches der beiden Flugzeuge befindet sich in einer geringeren Höhe?“ Die Beantwortung der Frage kann von der getesteten Person auf einen späteren Zeitpunkt verschoben werden (Endsley, 2019). Gemessen wird die RT bis Controller den Anruf annimmt als Indikator für WL und die Richtigkeit der Antwort als Indikator für SA (Jeannot et al., 2003).</p>
Vorbereitung	Anforderungsanalyse für SA, Fragen formulieren
Validierung	<p>Loft et al. (2015): Untersuchung verschiedenster Testverfahren bei $N = 171$ Studierenden, die eine Unterwasser Tracking-Aufgabe durchführten:</p> <ul style="list-style-type: none"> <input type="checkbox"/> SPAM korreliert mit ATWIT und NASA-TLX, aber nicht mit SART <input type="checkbox"/> SPAM besitzt prädiktive Validität <p>Durso et al. (2006): Untersuchung von $N = 88$ Personen, die verschiedene kognitive Testverfahren und SPAM bearbeiteten, während oder nachdem sie eine Flugsimulationsaufgabe durchführten:</p> <ul style="list-style-type: none"> <input type="checkbox"/> SPAM besitzt prädiktive Validität

	<ul style="list-style-type: none"> <input type="checkbox"/> SPAM als Verfahren zur Messung von SA besitzt inkrementelle Validität (von 9 %) zusätzlich zu einer Reihe kognitiver Tests (z.B. Raven-Matrizen, Tests der Federal Aviation Administration)
Bewertung	<p>Vorteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Einfaches und schnelles Verfahren, <input type="checkbox"/> vielversprechende Resultate in Validierungsstudien (Salmon et al., 2006) <input type="checkbox"/> gute Prädiktion von Performanz in Vielzahl von Studien (Endsley, 2019) <input type="checkbox"/> keine <i>task freezes</i> nötig: kann auch im Feld angewendet werden, <input type="checkbox"/> als Alternative zum SAVANT empfohlen (Langan-Fox et al., 2009). <p>Nachteile:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Aufwendige Vorbereitung, <input type="checkbox"/> geringe Konstruktvalidität (Salmon et al., 2006) <input type="checkbox"/> Intrusivität der Befragung: Negativer Einfluss auf Performanz und Erhöhung der WL (Endsley, 2019) <input type="checkbox"/> Sampling Bias durch Aufschub der Befragung: Fragen werden v.a. bei geringer WL beantwortet (Endsley, 2019) <input type="checkbox"/> korrektere Antworten, aber auch längere RT, da alle Informationen in der Simulation gesucht werden (Endsley, 2019) <input type="checkbox"/> Aufmerksamkeit wird als Folge des Messverfahrens evtl. nur auf aufgabenrelevante Aspekte gelenkt. <p>➔ Interessanter Ansatz, die SA über Wissen um zukünftige Ereignisse zu messen und Geschwindigkeit des Informationsabrufs und -verarbeitung als wesentlichen Aspekt von SA und guter Performanz zu definieren; aber verglichen mit anderen Verfahren (z.B. SAGAT und SART) diagnostisch schwächer.</p>
Verfügbarkeit	

3.1.15 Temporal Awareness

Autoren	Grosjean und Terrier (1999)
Verfahren	Objektiv
Datenerhebung	<i>Performance measure</i> , automatische computerbasierte Erfassung jeder Aktivität der Teilnehmer; Gewichtung der Daten notwendig, um diese vergleichen zu können
Durchführung	<p>Annahme: Einer Handlung geht das (un)bewusste Abwägen verschiedener Handlungsalternativen und das Priorisieren möglicher Handlungsschritte voraus, das in einer Strategie oder auch <i>cognitive trade-off</i> (Grosjean & Terrier, 1999) resultiert. Wenn Fehler begangen wurden, sollten die jeweiligen <i>trade-offs</i> untersucht werden, um den Handlungsprozess zu optimieren. Ein wichtiger Bestandteil der <i>cognitive trade-offs</i> ist Temporal Awareness.</p> <p>Temporal Awareness wird definiert als Bewusstsein um kürzlich vergangene und zukünftige Situationen, zusätzlich zum Bewusstsein über die aktuelle Situation. Laut Grosjean und Terrier (1999) ist Temporal Awareness wichtig für <i>process management tasks</i>, also besonders relevant bei mehreren zeitgleich ablaufenden Prozessen, die zu unterschiedlichen Zeitpunkten kontrolliert oder beeinflusst werden müssen.</p> <p>Temporal Awareness wird gemessen über die Anzahl von Fehlern bezüglich Zeit und Reihenfolge während eines Produktionsprozesses, Anzahl der Zeitperioden mit zeitlichen Hemmnissen, und relevante Zeitpunkte (landmarks), die von einem Operator berichtet werden.</p>
Validierung	Nach derzeitigem Wissensstand liegen keine Validierungsstudien vor.
Bewertung	<p>Temporal Awareness korreliert positiv mit einer geringeren Fehlerrate und somit einer besseren Leistung, was aber nicht generalisierbar ist, da dies nur für manche Fehlerarten zutrifft (Grosjean & Terrier, 1999).</p> <p>Interessantes Konstrukt, aber empirisch (noch) nicht überzeugend untersucht.</p>
Verfügbarkeit	

3.1.16 VESARS

	Virtual Environment Situation Awareness Rating System
Autoren	Scielzo et al. (2010)
Verfahren	Multidimensional, subjektives und objektives Verfahren
Datenerhebung	Performance measure, Szenarien mit einer Dauer von ca. 15 – 20 Minuten; Einschätzung der SA durch Expertinnen (Scielzo et al., 2010), drei Dimensionen, in virtueller und realer Umgebung durchführbar, Einzel- und Gruppentestungen
Durchführung	Messung von SA in virtuellen Trainingsumgebungen, die reale Orte darstellen sollen. Subjektive und objektive Daten werden in drei Bereichen (S. Abb. 25). erhoben, um individuelle SA, geteilte SA und Team-SA zu quantifizieren. Die Daten können wiederum in Echtzeit von Instruktor*innen genutzt werden, um individuelles Feedback während oder nach einer Trainingssession zu geben, und Auszubildende so unmittelbarer zu schulen. (Scielzo et al., 2010; Strater et al., 2013). Das VESARS wurde in das virtuelle Trainingsgame <i>Virtual Battlespace 2</i> integriert (Strater et al., 2013).
Vorbereitung	Voraussetzungen: Die Rater sollten gut geschult und vertraut mit dem Rating-system sei; die Fragen müssen vorab präzise formuliert werden (Gawron, 2019).
Validierung	Nach derzeitigem Wissensstand liegen keine Validierungsstudien vor.
Bewertung	Sehr interessanter Ansatz, aber weitere Validierungsstudien fehlen und die Anwendung ist auf ein schmales Feld beschränkt.
Verfügbarkeit	Scielzo et al. (2010)

Real-time SA queries	<ul style="list-style-type: none">• Where is the enemy?• What is your current direction of movement?• Will you be exposed on the route?
SA behaviors	<ul style="list-style-type: none">• Scans area, avenues of approach, and along routes• Tactically disperses to gather information• Uses tools and equipment (flashlight, etc.)
SA communications	<ul style="list-style-type: none">• Communication to prevent errors in performance• Statements regarding unit status• Attempting to prioritize task accurately

Abbildung 25

Beschreibung der drei Dimensionen von VESARS, Abbildung aus Scielzo et al. (2010)

3.2 Planungshilfen und Empfehlungen zu SA-Erfassung

Anders als der NASA-TLX bei WL, gibt es den klassischen SA-Test nicht. Salmon et al. (2006) beschreiben allerdings eine nützliche Übersicht und Bewertung verschiedenster SA Testverfahren. Diese zu konsultieren wird sich als besonders hilfreich erweisen, wenn es sich um die Messung von SA in sog. C4_i – Umgebungen dreht, also um hochtechnisierte Systeme, in denen die vier C (*command, control, communication, computers*) im Sinne maschineller Intelligenz von Bedeutung sind. Etwas allgemeiner auf alle HF-Situationen ausgelegt haben Endsley und Garland (2008) eine verdienstvolle und ebenfalls sehr ausführliche Aufstellung vorgelegt, in der sie auf über 300 Seiten die theoretischen Aspekte von SA beschreiben, sowie verschiedenste Verfahren (ein ganzes Kapitel über SAGAT), eine Meta-Analyse über 104 SA-Testverfahren und viele weitere Informationen rund um die Messung von SA anbieten. Ergebnisse der Meta-Analyse beschreiben unter anderem, dass Leistungsmessungen und subjektive Ratings generell sensitiver sind als Einschätzungen der SA nach Aufgaben (sog. *memory probes*).

Kürzlich hat Mica Endsley eine Übersicht (siehe Abb. 26) vorgelegt, die als eine schnelle Entscheidungshilfe dienen kann, wenn die Benutzung von SAGAT oder SPAM erwogen wird. Sie gruppiert sehr übersichtlich in indirekte Prozess- und Performanzmaße sowie direkte subjektive und objektive SA-Maße, die sich gut auf die Bewertung anderer Messverfahren von SA übertragen lassen.

TABLE 1: Comparison of SA Measurement Approaches

Process Measures	Performance Measures	Direct SA Measures		
		Subjective		Objective
Metrics Eye Tracking, Communications, Verbal Protocols, Physiological	Response time, Errors	Likert-type scales, SART	SAGAT	SPAM
Advantages <ul style="list-style-type: none"> Objective and continuous Eye tracking provides information on order and duration of attention to visual information Communications and verbal protocols can provide information on processes, strategies, types of assessments made 	<ul style="list-style-type: none"> Objective or subjective Can be gathered without operator input Often already collected 	<ul style="list-style-type: none"> Easy to collect Can be used across many domains Provides indication of confidence in SA 	<ul style="list-style-type: none"> Queries people on relevant SA knowledge on perceptions, comprehension, and projection Objectively scored based on simulation data Unbiased sampling across scenario avoids end-of-trial memory dependence 	<ul style="list-style-type: none"> Queries people on relevant SA knowledge of past, present, and future Objectively scored and timed based on simulation data Simulation freeze not required
Disadvantages <ul style="list-style-type: none"> Eye tracking does not capture attention to auditory cues or if information is correctly understood or integrated for higher levels of SA Communications provides only partial information on what is attended to and how processed. Some people verbalize more than others. Little research to date to support validity of physiological measures for SA 	<ul style="list-style-type: none"> Assumes what behavior will occur given a particular state of SA. System or training changes may affect performance in unexpected ways. SA for normal and emergency events may be different, so inferences constrained by scenarios tested and performance measures collected Confuses SA with performance which can be affected by other factors. Often insufficient sensitivity and diagnosticity. 	<ul style="list-style-type: none"> People may not be aware of what they do not know; meta-awareness is poor May be overly influenced by self assessments of performance Some scales (e.g., SART) include measures of workload 	<ul style="list-style-type: none"> Requires freezing of simulation scenario Requires people to answer queries based on memory for 2-3 min during freeze Requires development of domain-specific queries 	<ul style="list-style-type: none"> Requires dual-tasking to answer queries while performing task, potentially interfering with performance and creating a secondary task workload measure Allows people to "look-up" answers to queries which may not assess SA Requires development of domain-specific queries

Note. SA = situation awareness; SART = Situational Awareness Rating Technique; SAGAT = Situation Awareness Global Assessment Technique; SPAM = Situation Present Assessment Technique.

Abbildung 26
Vergleich von SAGAT und SPAM aus Endsley (2021)

4 Testverfahren zur Erfassung von Stress

4.1 Bewertungsprozesse/ Stresswahrnehmung

4.1.1 ABF

ABF (Alltags-Belastungs-Fragebogen)	
Autoren	Traue, Hrabal und Korarz (2000)
Inhalt	<ul style="list-style-type: none"> - Übersetzung des <i>Daily Stress Inventory</i> (Brantley, Waggoner, Jones & Rappaport, 1987) - Selbstrating - Instrument zur Erfassung von subjektiv belastenden Alltagsereignissen (der letzten 24 h) - Eignet sich für Zusammenhänge zwischen Alltagsstress – psychosomatischer/chronischer Erkrankungen
Theorie	<ul style="list-style-type: none"> - Theorie der Alltagsbelastungen (<i>daily hassles</i>) entwickelt aus Theorie der kritischen Lebensereignisse (<i>life events</i>)
Merkmale	<ul style="list-style-type: none"> - <i>Paper & pencil/online</i>, 61 Items, 7-stufig - Deutsch ➔ Frequenz = Anzahl der aufgetretenen Ereignisse ➔ Summe = Summe der Bewertungen der Ereignisse ➔ DB = Durchschnittliche Belastung = Summe/Frequenz
Durchführung	
Validierung	<p>Validierung (Traue et al., 2000):</p> <ul style="list-style-type: none"> - Gesunde Probanden (N=115), psychosomatisch erkrankte Patienten (N=451) - Retest-Reliabilität: $r = .36$ ➔ Es wird darauf hingewiesen, dass eine hohe Retest-Reliabilität nicht erwünscht ist, aufgrund des hohen Fluktuationspotentials von Tag zu Tag (24h Messung) - Konvergente Validität: geringe, aber signifikante Korrelation mit emotionaler Labilität - Normierung nicht vorhanden

Bewertung	- Ein geeignetes Messinstrument zur Erfassung von alltäglichen Belastungen (Traue et al., 2000)
Verfügbarkeit	Fragebogen und Validierung frei verfügbar in (Traue et al., 2000)

1. Ich habe eine Aufgabe nicht gut bewältigt. 0 1 2 3 4 5 6 7
2. Wegen anderer Personen habe ich schlechte Leistungen erbracht. 0 1 2 3 4 5 6 7
3. Ich dachte an meine unerledigten Arbeiten. 0 1 2 3 4 5 6 7
4. Ich musste hetzen, um einen Termin einzuhalten. 0 1 2 3 4 5 6 7
5. Ich wurde bei einer Aufgabe oder Arbeit gestört. 0 1 2 3 4 5 6 7
6. Jemand hat meine schon erledigte Arbeit mies gemacht. 0 1 2 3 4 5 6 7
7. Ich musste etwas machen, was ich nicht richtig konnte. 0 1 2 3 4 5 6 7
8. Ich konnte eine Aufgabe nicht zu Ende führen. 0 1 2 3 4 5 6 7
9. Ich war durcheinander. 0 1 2 3 4 5 6 7
10. Ich wurde kritisiert oder beschimpft. 0 1 2 3 4 5 6 7
11. Man hat mich nicht beachtet. 0 1 2 3 4 5 6 7
12. Ich musste öffentlich sprechen bzw. etwas vormachen. 0 1 2 3 4 5 6 7
13. Ich wurde unfreundlich bedient. 0 1 2 3 4 5 6 7
14. Man hat mich beim Sprechen unterbrochen. 0 1 2 3 4 5 6 7
15. Obwohl ich nicht wollte, musste ich mit Leuten zusammen sein. 0 1 2 3 4 5 6 7
16. Jemand hat ein Versprechen gebrochen/mich versetzt. 0 1 2 3 4 5 6 7
17. Ich stand in Konkurrenz zu jemandem. 0 1 2 3 4 5 6 7
18. Ich wurde angestarrt. 0 1 2 3 4 5 6 7
19. Jemand, an dem mir liegt, hat nichts von sich hören lassen. 0 1 2 3 4 5 6 7
20. Ich bin herumgeschubst worden. 0 1 2 3 4 5 6 7
21. Ich wurde missverstanden. 0 1 2 3 4 5 6 7
22. Man machte mich verlegen. 0 1 2 3 4 5 6 7
23. Mein Schlaf wurde gestört. 0 1 2 3 4 5 6 7
24. Ich habe etwas vergessen. 0 1 2 3 4 5 6 7
25. Ich hatte Angst vor Krankheit oder Schwangerschaft. 0 1 2 3 4 5 6 7
26. Mir ging es körperlich nicht gut bzw. ich war krank. 0 1 2 3 4 5 6 7
27. Jemand hat sich etwas ausgeliehen, ohne mich zu fragen. 0 1 2 3 4 5 6 7
28. Mir ist etwas kaputt gemacht worden. 0 1 2 3 4 5 6 7
29. Ich hatte einen kleinen Unfall (mir ist etwas zerbrochen, ein Kleidungsstück wurde zerrissen etc.). 0 1 2 3 4 5 6 7
30. Ich dachte an die Zukunft. 0 1 2 3 4 5 6 7
31. Lebensmittel oder Gegenstände meines persönlichen Bedarfs sind mir ausgegangen. 0 1 2 3 4 5 6 7
32. Ich hatte Streit mit meinem Partner/meiner Partnerin oder meinem Freund/meiner Freundin. 0 1 2 3 4 5 6 7
33. Ich hatte mit jemand anderem Streit. 0 1 2 3 4 5 6 7
34. Ich musste länger warten, als ich wollte. 0 1 2 3 4 5 6 7
35. Ich wurde beim Nachdenken oder Entspannen gestört. 0 1 2 3 4 5 6 7
36. Jemand hat sich vorgedrängt. 0 1 2 3 4 5 6 7
37. Ich war beim Sport/Spiel schlecht. 0 1 2 3 4 5 6 7
38. Ich habe etwas getan, was ich eigentlich nicht wollte. 0 1 2 3 4 5 6 7
39. Ich konnte nicht alles erledigen, was ich mir vorgenommen hatte. 0 1 2 3 4 5 6 7
40. Ich hatte mit dem Auto Probleme. 0 1 2 3 4 5 6 7

41. Der Verkehr hat mich nervös gemacht.	0 1 2 3 4 5 6 7
42. Ich hatte finanzielle Sorgen.	0 1 2 3 4 5 6 7
43. In einem Geschäft fand ich nicht das, was ich wollte.	0 1 2 3 4 5 6 7
44. Ich habe etwas verlegt.	0 1 2 3 4 5 6 7
45. Das Wetter war schlecht.	0 1 2 3 4 5 6 7
46. Ich hatte unerwartete Ausgaben (Bußgeld, Strafmandat).	0 1 2 3 4 5 6 7
47. Ich war mit einer Autoritätsperson konfrontiert.	0 1 2 3 4 5 6 7
48. Ich erhielt eine schlechte Nachricht.	0 1 2 3 4 5 6 7
49. Ich war um mein Äußeres besorgt.	0 1 2 3 4 5 6 7
50. Ich war einer bedrohlichen Situation ausgesetzt.	0 1 2 3 4 5 6 7
51. Ich habe mich über eine Fernsehsendung, einen Film, ein Buch geärgert.	0 1 2 3 4 5 6 7
52. Ich war schlecht gelaunt, weil ich gestört wurde (jemand hat nicht angeklopft, war unhöflich usw.).	0 1 2 3 4 5 6 7
53. Ich habe etwas nicht verstanden.	0 1 2 3 4 5 6 7
54. Ich habe mir Sorgen um jemanden gemacht.	0 1 2 3 4 5 6 7
55. Ich konnte gerade noch einer Gefahr ausweichen.	0 1 2 3 4 5 6 7
56. Ich habe eine schlechte Angewohnheit unterlassen (auf Nägel beißen, zu viel essen, Rauchen).	0 1 2 3 4 5 6 7
57. Ich hatte mit meinen Kindern Ärger.	0 1 2 3 4 5 6 7
58. Ich kam zur Arbeit/zu einer Verabredung zu spät.	0 1 2 3 4 5 6 7

Welche anderen unangenehmen Ereignisse sind heute aufgetreten?

59.
60.

**61. Hatten Sie heute ein besonders angenehmes Erlebnis?
Wenn ja, welches?**

Abbildung 27

Items des ABF (Traue et al., 2000)

4.1.2 TICS

TICS (Trierer Stressinventar)	
Autoren	Schulz, Schlotz und Becker (2004) (aktualisierte Version)
Inhalt	<ul style="list-style-type: none"> - Selbstrating , Ab 16 Jahren - mehrdimensional Variablen/Skalen: <ul style="list-style-type: none"> - Hohe Anforderungen: Arbeitsüberlastung, soziale Überlastung, Erfolgsdruck - Mangel an Bedürfnisbefriedigung: Arbeitsunzufriedenheit, Überforderung bei der Arbeit, Mangel an sozialer Anerkennung, soziale Spannungen, soziale Isolation - chronischer Stress - Screening-Kurzskala (SSCS): enthält die 12 aussagekräftigsten Items von den 5 Skalen: Chronische Besorgnis, Überforderung, Mangel an sozialer Anerkennung, arbeitsbezogene und soziale Überlastung - liefert ein Globalmaß für Stress
Theorie	Transaktionale Stressmodell (Stress ist Interaktion aus Umwelt und Ressourcen) Stress entsteht bei Dysbalance zwischen Stress und Ressourcen
Merkmale	<ul style="list-style-type: none"> - <i>Paper-pencil/online</i> - deutsch - Standardisiert: 57 Items, 5-stufig (wie oft eine bestimmte Situation erlebt wurde)
Durchführung	Dauer: 10-15 min (SSCS: 3 min)
Validierung	Zuverlässigkeit: Interne Konsistenz: $\alpha = .84-.91$ Profilreliabilität = .72 Testgültigkeit: Konstruktvalidität erfüllt: Faktorenanalysen, Korrelationen mit Stressfragebögen, Persönlichkeitsmerkmalen, Partnerschaftsverhalten, sozialer Unterstützung, Schlafqualität, körperlichen und psychischen Beschwerden, Cortisolausschüttung

	Normierung liegt vor (N=604): Gesamtstichprobe sowie drei Altersgruppen (16-30, 31-59, 60-70 Jahre).
Bewertung	sorgfältig entwickelt, normiert, langjährig erprobt, theoretisch gestützt, sehr gute Gütekriterien Einsatz in Forschung und Praxis (Diagnostik, Evaluation von Stresstrainings)
Verfügbarkeit	Schulz et al., 2004 (→ Hogrefe)

4.1.3 AVEM

Arbeitsbezogenes Verhaltens- und Erlebnismuster	
Autoren	Schaarschmidt & Fischer (2008) (aktualisierte Version)
Inhalt	<ul style="list-style-type: none"> - mehrdimensionales persönlichkeitsdiagnostisches Verfahren - gesundheitsförderliche bzw. -gefährdende Verhaltens- und Erlebensweisen bei der Bewältigung von Arbeits- und Berufsanforderungen (z. B. Risikomuster, die im Zusammenhang mit Burnout stehen) - Ergebnisse ermöglichen: Begründung und Ableitung präventiver Maßnahmen sowie deren Erfolgskontrolle im Rahmen der beruflichen Rehabilitation - 11 Dimensionen → 4 Erlebnismuster: G (Gesundheit), S (Schonung), A (Risiko i. S. der Selbstüberforderung) und B (Risiko i. S. von chronischem Erschöpfungserleben und Resignation) - für jede Person wird die Ähnlichkeit ihres Profils mit diesen vier Referenzmustern ermittelt
Theorie	
Merkmale	<ul style="list-style-type: none"> - Selbstrating - Deutsch - Paper-pencil/PC (wird mit computergestützten Auswertungsprogramm geliefert), 66 Items, 5-stufig - Kurzform, 44 Items - individuell und in Gruppen
Durchführung	Dauer: 12 min (Kurzform: 8 min)
Validierung	<p>Zuverlässigkeit:</p> <p>Hohe interne Konsistenz: $\alpha = .79 - .87$; Kurzform: $\alpha = .75 - .83$</p> <p>Testgültigkeit (N=31.979):</p> <p>Normen in Bezug auf die Skalenwerte (Berufsgruppen, Studierende/Auszubildende, Patienten)</p>
Bewertung	<p>Rüesch et al., 2006:</p> <ul style="list-style-type: none"> - Insbesondere zur Ableitung präventiver Maßnahmen geeignet - Aussagen über Einstellung, Befinden und Verhaltensänderung nach Interventionen
Verfügbarkeit	Schaarschmidt & Fischer, 2008 (→ Pearson)

4.1.4 IR

IR (Irritationsskala)	
Autoren	Mohr, Müller & Rigotti (2007)
Inhalt	<ul style="list-style-type: none"> - Screening-Verfahren für die betriebliche Gesundheitsanalyse Evaluation von stressreduzierenden Maßnahmen ableitbar - Irritation als Indikator psychischer (Fehl-)Beanspruchung in der Folge von Belastungen durch Arbeit und Vorläufer weiterer psychischer Befindensbeeinträchtigungen. - mehrdimensional <p>Unterscheidung:</p> <ul style="list-style-type: none"> - Kognitive Irritation (arbeitsbezogene Rumination → tätigkeitspezifische Stressoren) - Emotionale Irritation (Gereiztheit → soziale Stressoren)
Theorie	
Merkmale	<ul style="list-style-type: none"> - Englisch - Selbstrating, 8 Items, 7-stufig - Einzel- oder Gruppentest
Durchführung	Dauer: wenige Minuten
Validierung	<p>Zuverlässigkeit:</p> <p>hohe interne Konsistenz:</p> <p>emotionale Irritation: $\alpha = .80 - .90$</p> <p>kognitive Irritation zwischen $\alpha = .75$ und $\alpha = .91$.</p> <p>Die Retest-Reliabilität (Gesamtindex Irritation) beträgt nach 6 Monaten $r_{tt} = .69$, nach zweieinhalb Jahren $.61$, nach dreieinhalb Jahren $.57$.</p> <p>Kognitive Irritation weist eine geringere Stabilität auf als die Emotionale Irritation, da sie stärker mit den aktuell vorhandenen Belastungen korrespondiert.</p> <p>Testgültigkeit:</p> <p>Testzentrale:</p> <p>hoch valide in zahlreichen Studien</p> <p>Irritation als konvergent/diskriminant</p>

	Stanine-Normen (N= 4735)
Bewertung	(Rüesch et al., 2006): <ul style="list-style-type: none"> - Insbesondere sinnvoll zur Maßnahmenevaluation - Branchenübergreifendes Instrument - Sehr ökonomisch um potentiell schädigende Wirkungen kritischer Arbeitsbedingungen zu erfassen
Verfügbarkeit	Nicht kommerziell vertrieben: Anfragen an die Autoren (Rüesch et al., 2006)

Items

KI = Kognitive Irritation
EI = Emotionale Irritation

Nr.	Item	Sub- skala
1	Es fällt mir schwer, nach der Arbeit abzuschalten.	KI
2	Ich muss auch zu Hause an Schwierigkeiten bei der Arbeit denken.	KI
3	Wenn andere mich ansprechen, kommt es vor, dass ich mürrisch reagiere.	EI
4	Selbst im Urlaub muss ich manchmal an Probleme bei der Arbeit denken.	KI
5	Ich fühle mich ab und zu wie jemand, den man als Nervenbündel bezeichnet.	EI
6	Ich bin schnell verärgert.	EI
7	Ich reagiere gereizt, obwohl ich es gar nicht will.	EI
8	Wenn ich müde von der Arbeit nach Hause komme, bin ich ziemlich nervös.	EI

Abbildung 28

Items des IR (Mohr et al., 2007)

4.1.5 PSS

PSS (Perceived Stress Scale)	
Autoren	Deutsch: Schneider et al., (2020) Original (Englisch): Cohen et al., 1983)
Inhalt	<ul style="list-style-type: none"> - Erfassung des wahrgenommenen Stress - mehrdimensional - Skalen: Hilflosigkeit, Selbstwirksamkeit (& Gesamtscore)
Theorie	
Merkmale	Selbststrating, 10 Items, 5-stufig (1-5)
Durchführung	Dauer: wenige Minuten
Validierung	Validierungsstudie der deutschen Version (N=2463) (Klein et al., 2016 – mit deutscher Übersetzung) Zuverlässigkeit: Cronbachs Alpha $\alpha = .84$ Testgültigkeit: Konstruktvalidität (Faktorenanalysen) gut Normierung: Orientierung an Validierungsstudie (Klein et al., 2016)
Bewertung	Reliables, valides, ökonomisches Instrument zur Bewertung des Stressempfindens
Verfügbarkeit	Frei verfügbar

PERCEIVED STRESS SCALE - 10

Die folgenden Fragen beschäftigen sich mit Ihren Gedanken und Gefühlen während des letzten Monats. Bitte geben Sie für jede Frage an, wie oft sie in entsprechender Art und Weise gedacht oder gefühlt haben.

		nie	Fast nie	Manchmal	Ziemlich oft	sehr oft
1	Wie oft waren Sie im letzten Monat aufgewühlt, weil etwas unerwartet passiert ist?	1	2	3	4	5
2	Wie oft hatten Sie im letzten Monat das Gefühl, nicht in der Lage zu sein, die wichtigen Dinge in Ihrem Leben kontrollieren zu können?	1	2	3	4	5
3	Wie oft haben sie sich im letzten Monat nervös und gestresst gefühlt?	1	2	3	4	5
4	Wie oft waren Sie im letzten Monat zuversichtlich, dass Sie fähig sind, ihre persönlichen Probleme zu bewältigen?	1	2	3	4	5
5	Wie oft hatten Sie im letzten Monat das Gefühl, dass sich die Dinge zu Ihren Gunsten entwickeln?	1	2	3	4	5
6	Wie oft hatten Sie im letzten Monat den Eindruck, nicht all Ihren anstehenden Aufgaben gewachsen zu sein?	1	2	3	4	5
7	Wie oft waren Sie im letzten Monat in der Lage, ärgerliche Situationen in Ihrem Leben zu beeinflussen?	1	2	3	4	5
8	Wie oft hatten Sie im letzten Monat das Gefühl, alles im Griff zu haben?	1	2	3	4	5
9	Wie oft haben Sie sich im letzten Monat über Dinge geärgert, über die Sie keine Kontrolle hatten?	1	2	3	4	5
10	Wie oft hatten Sie im letzten Monat das Gefühl, dass sich so viele Schwierigkeiten angehäuft haben, dass Sie diese nicht überwinden konnten?	1	2	3	4	5

Skala Hilflosigkeit (H): Summe der Items 1, 2, 3, 6, 9, 10; Skala Selbstwirksamkeit (S): Summe der Items 4, 5, 7, 8. Für die Berechnung des Gesamtscores müssen die Items 4, 5, 7 und 8 der Selbstwirksamkeitsskala invertiert werden. Der Gesamtscore berechnet sich aus der Summe der Items der Hilflosigkeitsskala und der Summe der invertierten Items der Selbstwirksamkeitsskala. Höhere Werte deuten auf ein erhöhtes Stresslevel hin.

Abbildung 29

Items des PSS (Schneider et al., 2020)

4.1.6 Mental Health Fragebogen

Mental Health	
Autoren	Luftfahrt Bundesamt
Inhalt	Flugmedizinische Tauglichkeitsuntersuchung Ziviler Einsatz
Theorie	
Merkmale	Selbstrating, 20 Items, 5-stufig
Durchführung	
Validierung	Keine Angaben
Bewertung	Keine Angaben
Verfügbarkeit	https://www.fliegerdoc-muenchen.de/downloads/Fragebogen_Mental_Health.pdf

		Zutreffendes bitte ankreuzen! Mark with a cross where applicable!				
		Stimme gar nicht zu I fully disagree	Stimme eher nicht zu I tend to disagree	weder nicht neither agree nor disagree	Stimme eher zu I tend to agree	Stimme völlig zu I fully agree
1	Meine Schlafgewohnheiten haben sich nicht verändert. / My sleeping habits did not change.					
2	Ich hatte in letzter Zeit häufig starken Stress. Recently, I have often experienced high levels of stress.					
3	Ich fühle mich in meinem Beruf zu wenig/nicht herausgefordert. In my job I am hardly challenged.					
4	Ich kann die Erwartungen und Bedürfnisse meiner Familie nur schwer erfüllen. / It is difficult for me to meet the expectations and needs of my family.					
5	Ich habe Schuldgefühle, wenn ich Alkohol getrunken habe. I feel guilty after drinking alcohol.					
6	In letzter Zeit bin ich öfter gereizt. I often feel irritable recently.					
7	Meine finanzielle Situation ist angespannt. My financial situation is tight.					
8	Ich schlafe leicht ein und meistens die ganze Nacht durch I fall asleep easily and sleep through the whole night most nights					
9	Ich entspanne mich nach der Arbeit gerne bei einem Drink. I like to relax with a drink after a day of work.					

Self-Description Form in the Course of the Aeromedical Examination

		Zutreffendes bitte ankreuzen! Mark with a cross where applicable!				
		Stimme gar nicht zu I fully disagree	Stimme eher nicht zu I tend to disagree	weder nicht neither agree nor disagree	Stimme eher zu I tend to agree	Stimme völlig zu I fully agree
10	Es fällt mir in letzter Zeit schwer, meine Stimmung zu kontrollieren. Recently I have difficulties controlling my mood.					
11	Unter meinen Kollegen habe ich wenige oder keine Freunde. I do not have many (or any) friends at my work place.					
12	Mein Beziehungsstatus wechselt häufig. My relationship status changes frequently.					
13	Mein Gewicht schwankt häufig. My body weight varies frequently.					
14	Wir haben in der Firma kein gutes Arbeitsklima. The work climate in my company is not positive.					
15	Es fällt mir leicht, mein Selbstvertrauen aufrecht zu halten. I find it easy to maintain my self confidence.					
16	Ich brauche Medikamente oder andere Substanzen um meine Stimmung positiv zu halten. / I need medication or other substances to maintain a positive mood.					
17	Ich überlege derzeit einen Jobwechsel. I am currently considering a career change.					
18	Es fällt mir schwer, von meiner aktuellen Arbeit und Aufgaben abzuschalten. / Currently I find it difficult to take my mind off my work and tasks.					
19	Es fällt mir nach einigen Tagen Arbeit schwer, meinen Schlafrhythmus wieder zu normalisieren. / I find it difficult to regain my sleeping rhythm after several days at work.					
20	Ich bin überzeugt, dass Alkohol und andere stimmungsverändernde Substanzen die Leistungsfähigkeit nicht beeinträchtigen. / I am confident that alcohol and other mood altering substances do not effect my work capacity.					

Abbildung 30
Items des Mental Health Fragebogens (Luftfahrt-Bundesamt)

4.2 Stressbewältigung

4.2.1 RESTQ/EBF

Recovery stress questionnaire/Erholungs-Belastungs-Fragebogen	
Autoren	Kallus & Kellmann (2016) (aktualisierte Version)
Inhalt	<ul style="list-style-type: none"> - Gegenwärtiger Erholungs-Beanspruchungs-Zustand - Vollständiges Bild des aktuellen Ausmaßes von Belastung und Erholung - Mehrdimensional: Allgemeine Version (RESTQ/EBF-Basic): 7 Belastungsskalen und 5 Erholungsskalen - Spezifische Version: Sportler (RESTQ/EBF-Sport), für Trainer (RESTQ/EBF-Coach), für Kinder und Jugendliche (RESTQ/EBF-CA), für den Arbeitskontext (RESTQ/EBF-Work) und für den Einsatz im klinischen Bereich.
Theorie	<ul style="list-style-type: none"> - Annahme, dass eine Akkumulation von Belastungen in verschiedenen Lebensbereichen mit unzureichenden Möglichkeiten zur Erholung zu einem kritischen psychophysiologischen Zustand führt.
Merkmale	<p>Selbststrating, <i>Paper & pencil/online</i>, 7-stufig (0-6), Itemanzahl variiert</p> <p>Alter: 10-88 Jahre (je nach Version)</p> <p>Jede Version hat spezifisches Zeitfenster (3, 7, 14 Tagen/Nächten)</p>
Durchführung	Dauer: je nach Version (5-15 min)
Validierung	<p>Zuverlässigkeit:</p> <ul style="list-style-type: none"> - Cronbachs Apha: gut-zufriedenstellend - Retest-Koeffizient: Test-Retest-Koeffizienten sind bei der Messung in kurzen zeitlichen Abständen hoch, nehmen bei Messungen mit längeren Zeitabschnitten ab (→im Sinne der Messintention!) <p>Testgültigkeit:</p> <ul style="list-style-type: none"> - Konstruktvalidität erfüllt: Interkorrelationen der Skalen, die Faktorenstruktur und die Stabilität der Interkorrelationen über verschiedene Stichproben hinweg - Validierung des Erholungs-Beanspruchungs-Zustands: Vergleiche mit Stimmungszustand, Leistung, biologischen Indikatoren.
Bewertung	Geeignetes Instrument für Zustands- und Verlaufsdiagnostik im betrieblichen und klinischen Setting: entsprechend validiert und normiert (Rüesch et al., 2006)
Verfügbarkeit	Kallus & Kellmann, 2016 (→ Pearson)

4.2.2 CISS

CISS (Coping Inventar zum Umgang mit Stress-Situationen)	
Autoren	Deutsch: Kälin & Semmer (2020) Original: Endler & Parker (1999)
Inhalt	<ul style="list-style-type: none"> - Erfassung der habituellen Stressbewältigung (Coping) - mehrdimensional - Erhebung der 3 grundlegenden Coping-Stile: - Aufgabenorientiertes Coping: Versuche der direkten Problemlösung bzw. der Planung von geeigneten Schritten gegen die stressverursachenden Bedingungen - Emotionsorientiertes Coping: Auf die eigene Person bezogene, emotionale Reaktionen - Vermeidungsorientiertes Coping: Vermeidung der Auseinandersetzung mit Stress-Situationen. Das Vermeidungsorientierte Coping kann weiter unterteilt werden in <ul style="list-style-type: none"> a) Sozial-ablenkungsorientiertes Coping: Ablenkung von Problem-Situationen mittels Suche von sozialen Kontakten und b) Zerstreuungsorientiertes Coping: Allgemeine Ablenkung von Problem-Situationen
Theorie	-
Merkmale	<ul style="list-style-type: none"> - Selbstrating, 24 Items - Ab 16 Jahren
Durchführung	- Dauer: 10-15 min
Validierung	Zuverlässigkeit: <ul style="list-style-type: none"> - N=685 - Aufgabenorientiertes Coping: $\alpha = .83$ - Emotionsorientiertes Coping: $\alpha = .80$ - Vermeidungsorientiertes Coping: $\alpha = .79$ - Subskalen Sozial-ablenkungsorientiertes Coping: $\alpha = .80$ - Zerstreuungsorientiertes Coping: $\alpha = .71$

	<p>Testgültigkeit:</p> <ul style="list-style-type: none"> - Testwiederholungsreliabilitäten (N = 359) im Abstand eines Jahres: $r = .47$ (Zerstreuungsorientiertes Coping) und $r = .65$ (Emotionsorientiertes Coping) <p>Normstichprobe: N=665 (berufstätige Schweizer Erwachsene; Prozentränge und T-Werte)</p>
Bewertung	<ul style="list-style-type: none"> - Valide und reliabel - Anwendung vor allem: in Bereichen der Arbeits- und Organisationspsychologie, der Persönlichkeitspsychologie und in der Forschung bei gesundheitspsychologischen Fragestellungen.
Verfügbarkeit	Kälin & Semmer, 2020 (→ Hogrefe)

4.2.3 UBV

UBV (Fragebogen zum Umgang mit Belastungen im Verlauf)	
Autoren	Reicherts & Perrez (1993)
Inhalt	<ul style="list-style-type: none"> - Fragebogen zum Umgang mit Belastungen im Verlauf - Erfassung von psychologischen Merkmalen der Belastungsverarbeitung - Vorgegeben: verschiedene alltagsnahe Problemsituationen, jeweils in einer Sequenz von drei Phasen: dem Beginn, dem Fortbestehen und dem (positiven und negativen) Ausgang der Stressepisode - mehrdimensional - Variablen: Situationseinschätzungen, emotionale Reaktionen, Bewältigungsintentionen und kognitives, emotionales und instrumentelles Bewältigungsverhalten
Theorie	Lazarus & Folkman (1987)
Merkmale	<ul style="list-style-type: none"> - Deutsch - Alter: 20-65 - Selbstrating, 47 Items
Durchführung	<ul style="list-style-type: none"> - Dauer: Gesamtform (18 Belastungsepisoden) etwa 3 Stunden - Kurzform (4 Episoden) etwa 40 Minuten
Validierung	<p>Zuverlässigkeit:</p> <ul style="list-style-type: none"> - Interne Konsistenz: $\alpha=.71-.95$ - Retest-Reliabilität: $r = .80$ <p>Testgültigkeit:</p> <ul style="list-style-type: none"> - multiple Korrelation der UBV-Variablen: - mit Depression (BDI): $r = .80$, - mit Ängstlichkeit (STAI) über $r = .87$. <p>➔ Hohe Effektstärken in Interventionsstudien belegen die Treatment-Sensibilität verschiedener UBV-Variablen.</p> <p>Normen für die einzelnen Verlaufsstadien: Stanine; Prozentränge für nichtstudentische Erwachsene (N=100), Studenten (N=104)</p>
Bewertung	<ul style="list-style-type: none"> - validiert, Einsatz in Forschung und individualdiagnostischen Untersuchungen
Verfügbarkeit	Reicherts & Perrez, 1993 (➔ Verlag Hans Huber)

Reicherts/Perrez, Zum Umgang mit Belastungen

Tabelle 2: Übersicht über die Reaktions-Komponenten und Variablen des UBV

Komponente / Variable	Phase im Episodenverlauf			Belastungsart	
	1	2	3	aversiv	Verlust
<i>(1_) Emotionale Streßreaktionen</i>					
(11) Ängstlichkeit	x	x	x	x	x
(12) Deprimiertheit	x	x	x	x	x
(13) Aggressivität	x	x	x	x	x
(11,12,13) Negative Emotionalität	x	x	x	x	x
<i>(2_) Situationseinschätzungen</i>					
(21) Wandelbarkeit	x	x		x	x
(22) Kontrollierbarkeit	x	x		x	x
(23) negative Valenz	x	x	x	x	x
(24) Familiarität	x			x	x
<i>(3_) Bewältigungsintentionen</i>					
(31) Stressor beeinflussen	x	x		x	
(32) Konflikt meiden (mit Stressor)	x	x		x	
(33) Verlust abwenden	x	x			x
(34) Verlust substituieren	x	x			x
(35) Emotionale Equilibrierung	x	x		x	x
(36) Selbstwert-Equilibrierung	x	x		x	x
<i>(4_) Bewältigung selbstbezogen</i>					
(4 1) Informationsunterdrückung	x	x	x	x	x
(42) Palliation	x	x	x	x	x
(43) Informationssuche	x	x	x	x	x
(44) Umbewertung	x	x	x	x	x
(45) Fremdbeschuldigung	x	x		x	x
(46) Selbstbeschuldigung	x	x		x	x
(47) Selbstbekräftigung			x	x	x
(48) Soziale Unterstützung			x	x	x
<i>(5_) Bewältigung umgebungsbezogen</i>					
(51) Passivität (aversive Situationen)	x	x		x	
(52) Evasion	x	x		x	
(53) aktive Stressorbeeinflussung	x	x		x	
(54) aktives Verhindern	x	x			x
(55) Passivität (Verlustsituationen)	x	x			x
(56) aktives Umorientieren	x	x			x
<i>(6_) Kausalattributionen</i>					
(61) intern: eigenes Verhalten			x	x	x
(62) extern: Verhalten anderer			x	x	x
(63) extern: Umstände			x	x	x
<i>(7_) Palliationsinventar</i>					
„aktive Entspannung“ (7 1,73,74)	unabhängig von den Streßepisoden erfaßt				
„Entlastung“ (72,75,78)					
„Alltagsdrogen“ (76,77)					

Abbildung 31

Reaktionskomponenten und Variablen UBV (Reicherts & Perrez, 1993)

4.2.4 FERUS

FERUS (Fragebogen zur Erfassung von Ressourcen und Selbstmanagementfähigkeiten)	
Autoren	Jack (2007)
Inhalt	<ul style="list-style-type: none"> - Erfasst gesundheitsrelevante Ressourcen und Selbstmanagementfähigkeiten - mehrdimensional - 7 Skalen: Veränderungsmotivation, Selbstbeobachtung, aktives und passives Coping, Selbstwirksamkeit, Selbstverbalisation, Hoffnung und soziale Unterstützung
Theorie	<p>Salutogenese-Modell (Antonovsky, 1979)</p> <p>Selbstmanagement-Therapie (Kanfer, Reinecker & Schmelzer, 1996)</p> <p>Theorie der Selbstwirksamkeit nach Bandura (1977)</p> <p>Selbstinstruktionstechniken nach Meichenbaum (1977)</p> <p>Depressionstheorie (Beck, Steer, & Brown, 1987)</p> <p>soziale Unterstützungstheorie von Sommer und Fydrich (1989).</p>
Merkmale	<ul style="list-style-type: none"> - Selbstrating, 66 Items, 5-stufig - Einzel oder Gruppentestung
Durchführung	<ul style="list-style-type: none"> - Dauer: 20-30 min
Validierung	<p>Zuverlässigkeit:</p> <p>befriedigende bis gute Retest-Reliabilitäten (.66 bis .86)</p> <p>gute bis sehr gute interne Konsistenz (.86 bis .93)</p> <p>Testgültigkeit:</p> <p>→ Skalen sind inhaltlich valide und erfüllen die faktorielle sowie die kriterienbezogene Validität</p>
Bewertung	<p>Grundgedanke: Erwachsene in ambulanter oder stationärer psychotherapeutischer Behandlung.</p> <p>Einsatz bei Beratungsklienten, nicht akut psychotischen Psychatriepatienten und Mitarbeitern im Rahmen einer betrieblichen Gesundheitsförderung denkbar</p>
Verfügbarkeit	Jack, 2007 (→ Hogrefe)

4.2.5 SVF

SVF (Stressverarbeitungsfragebogen)	
Autoren	Erdmann & Janke (2008)
Inhalt	<ul style="list-style-type: none"> - Erfassung von Bewältigungs- bzw. Verarbeitungsmaßnahmen in belastenden Situationen - kein singuläres Testinstrument, sondern ein Inventar von Methoden - mehrdimensional - 20 Stressverarbeitungsweisen im Sinne zeit- und situations-(stressor-)stabiler Personenmerkmale: Bagatellisierung, Herunterspielen, Schuldabwehr, Ablenkung, Ersatzbefriedigung, Selbstbestätigung, Situationskontrolle, Reaktionskontrolle, Positive Selbstinstruktion, Soziales Unterstützungsbedürfnis, Vermeidung, Flucht, Soziale Abkapselung, Gedankliche Weiterbeschäftigung, Resignation, Selbstbemitleidung, Selbstbeschuldigung, Aggression, Pharmakaeinnahme, Entspannung. - Kurzform (13 Subtests)
Theorie	
Merkmale	<ul style="list-style-type: none"> - Selbstrating - Deutsch - Alter: 20-79 Jahre
Durchführung	Dauer: 10-15 min
Validierung	<p>Zuverlässigkeit: interne Konsistenz, zwischen: $\alpha = .66$ und $\alpha = .92$ Retest, zwischen: $r_{tt} = .69$ und $r_{tt} = .86$</p> <p>Testgültigkeit: Überprüfung von: Konstruktvalidität, differenzielle, kriterienbezogene Vorhersagevalidität</p> <p>Normierung vorhanden (T-Werte)</p>
Bewertung	Stressforschung (insbesondere klinisch-psychologische Forschung)
Verfügbarkeit	Erdmann & Janke, 2008 (→ Hogrefe)

4.2.6 SCS/D

SCS-D (Self-Compassion Scale)	
Autoren	Neff (2003) Deutsche Version: (Hupfeld & Ruffieux, 2011)
Inhalt	Self-Compassion = positive Grundeinstellung gegenüber der eigenen Person in schwierigen Lebenssituationen (emotionale Resilienz-Faktor) <ul style="list-style-type: none"> - mehrdimensional Faktoren: <ul style="list-style-type: none"> → Selbstbezogene Freundlichkeit → Selbstverurteilung → Verbindende Humanität → Isolation → Achtsamkeit → Überidentifizierung
Theorie	
Merkmale	<ul style="list-style-type: none"> - Selbstrating, 26 Items, 5-stufig - deutsch
Durchführung	
Validierung	Validierung des deutschen FB (Hupfeld & Ruffieux, 2011): Stichprobe 1: N= 396, Stichprobe 2: N= 165 Zuverlässigkeit: Cronbachs Alpha $\alpha > .90$ Retest: $r_{tt} = .93$ Testgültigkeit: Konstruktvalidität vorhanden: Erwartungsgemäße Zusammenhänge mit psychischer Belastung und subjektivem Wohlbefinden
Bewertung	Der SCS in deutscher Version ist ein geeignetes Instrument, mit dem Self-Compassion ökonomisch, reliabel und valide erfasst werden kann.
Verfügbarkeit	deutsche Version verfügbar in: Hupfeld & Ruffieux (2011)

Items der deutschsprachigen Version der Self-Compassion Scale (SCS-D)

1. Ich missbillige und verurteile meine eigenen Fehler und Schwächen. (SV)
2. Wenn ich mich niedergeschlagen fühle, neige ich dazu nur noch auf das zu achten, was nicht in Ordnung ist. (ÜI)
3. Wenn die Dinge bei mir schlecht laufen, sehe ich diese Schwierigkeiten als Teil des Lebens, den jeder einmal durchlebt. (VH)
4. Wenn ich über meine Fehler und Mängel nachdenke, neige ich dazu mich vom Rest der Welt getrennt und abgeschnitten zu fühlen. (I)
5. Ich versuche mit mir selbst liebevoll umzugehen, wenn es mir emotional schlecht geht. (SF)
6. Wenn ich bei etwas versage, was mir wichtig ist, werde ich von Gefühlen der Unzulänglichkeit aufgezehrt. (ÜI)
7. Wenn ich völlig am Ende bin, rufe ich mir in Erinnerung, dass es vielen anderen Menschen auf der Welt genauso geht. (VH)
8. In wirklich schwierigen Zeiten neige ich dazu, streng mit mir selbst zu sein. (SV)
9. Wenn mich etwas aufregt, versuche ich meine Gefühle im Gleichgewicht zu halten. (A)
10. Wenn ich mich auf irgendeine Art unzulänglich fühle, versuche ich mich daran zu erinnern, dass die meisten Leute solche Gefühle der Unzulänglichkeit haben. (VH)
11. Ich bin intolerant und unduldsam gegenüber denjenigen Seiten meiner Persönlichkeit, die ich nicht mag. (SV)
12. Wenn ich eine sehr schwere Zeit durchmache, schenke ich mir selbst die Zuwendung und Einfühlsamkeit, die ich brauche. (SF)
13. Wenn es mir schlecht geht, neige ich dazu zu glauben, dass die meisten anderen Menschen wahrscheinlich glücklicher sind als ich. (I)
14. Wenn etwas Unangenehmes passiert, versuche ich einen ausgewogenen Überblick über die Situation zu erlangen. (A)
15. Ich versuche, meine Fehler als Teil der menschlichen Natur zu sehen. (VH)
16. Wenn ich Eigenschaften bei mir feststelle, die ich nicht mag, dann deprimiert mich das. (SV)
17. Wenn ich bei etwas scheitere, das mir wichtig ist, versuche ich die Dinge nüchtern zu betrachten. (A)
18. Wenn ich wirklich zu kämpfen habe, neige ich zur Ansicht, dass andere es sicherlich einfacher haben. (I)
19. Ich gehe freundlich mit mir um, wenn ich Kummer und Leid erfahre. (SF)
20. Wenn mich etwas aufregt, werde ich von meinen Gefühlen förmlich mitgerissen. (ÜI)
21. Wenn ich Leid erfahre, kann ich mir gegenüber ein wenig kaltherzig sein. (SV)
22. Wenn es mir schlecht geht, versuche ich meinen Gefühlen mit Neugierde und Offenheit zu begegnen. (A)
23. Ich akzeptiere meine Fehler und Schwächen. (SF)
24. Wenn etwas Unangenehmes passiert, neige ich dazu, den Vorfall völlig zu übertreiben. (ÜI)
25. Wenn mir etwas für mich Wichtiges misslingt, glaube ich oft, dass nur ich allein versage. (I)
26. Ich versuche verständnisvoll und geduldig gegenüber jenen Zügen meiner Persönlichkeit zu sein, die ich nicht mag. (SF)

Anmerkungen: (SF) = Selbstbezogene Freundlichkeit, (SV) = Selbstverurteilung, (VH) = Verbindende Humanität, (I) = Isolation, (A) = Achtsamkeit, (ÜI) = Überidentifizierung; Antwortskala: 1 = *sehr selten*, 2 = *selten*, 3 = *gelegentlich*, 4 = *oft*; 5 = *sehr oft*.

Abbildung 32

Items SCS-D (Quelle: Hupfeld & Ruffieux (2011), Anhang)

4.3 Stresserleben

4.3.1 BELA

BELA (Belastungsfragebogen)	
Autoren	Boucsein (1973)
Inhalt	<ul style="list-style-type: none"> - Dient der Erfassung von Stressreagibilität - Proband antwortet auf 11 vorgegebene Belastungsszenarien: Angeben der Intensität seiner „innerlichen Reaktion“ - mehrdimensional - Belastungssituationen, z.B.: Schmerzreize, Konfrontation mit körperlichen Verletzungen, Versagen in sozialen Situationen, überhöhte Arbeitsanforderungen
Theorie	Konstrukt Stressreaktivität → Disposition einer Person, auf Belastungen mit entsprechender Stressreaktion zu antworten (entsprechend intensiv/lang/schnell)
Merkmale	Selbstrating, 11 Items, 5-stufig
Durchführung	Keine Angaben
Validierung	Keine Angaben
Bewertung	BELA als Vorläufer zur Erfassung von Coping Forschungseinsatz
Verfügbarkeit	Keine Angaben

4.3.2 SRS

SRS (Stress-Reaktivitäts-Skala)	
Autoren	Schulz, Jansen & Schlotz (2005)
Inhalt	<p>Erhebung der generellen und belastungsspezifischen Stressreaktivität</p> <ul style="list-style-type: none"> - mehrdimensional <p>Insgesamt 6 Skalen:</p> <ul style="list-style-type: none"> - Reaktivität auf Arbeitsüberlastung (Gefühl von Nervosität, Aufregung, Gereiztheit als Reaktion auf hohe Arbeitsbelastung) - Reaktivität auf soziale Konflikte (Betroffenheit, Verärgerung, Aufregung als Reaktion auf soziale Konflikte, Kritik, Zurückweisung) - Reaktivität auf sozialen Stress (nervös sein, Selbstvertrauen verlieren Selbstvertrauen als Reaktion auf soziale Bewertung) - Reaktivität auf Versagen (sich ärgern, enttäuscht sein, niedergeschlagen sein als Reaktion auf Versagen) - Antizipatorische Reaktivität (sich nervös fühlen, angespannt sein vor einer einer wichtigen Aufgabe oder Entscheidung) - verlängerte Reaktivität (Schwierigkeiten sich nach einer hohen Arbeitsbelastung zu entspannen/abzuschalten).
Theorie	<p>Konstrukt Stressreaktivität → Disposition einer Person, auf Belastungen mit entsprechender Stressreaktion zu antworten (entsprechend intensiv/lang/schnell)</p> <p>4 Determinanten als Personenmerkmale (für eine ungünstige Ausprägung):</p> <ul style="list-style-type: none"> - hohe Besorgnisneigung - negatives Selbstkonzept eigener Fähigkeiten - hohe Erregbarkeit des ZNS - hohe Negative Affektivität.
Merkmale	<ul style="list-style-type: none"> - deutsch - Selbstrating, 29 Items 3-stufig - Gesamtwert durch Aggregation der Teilskalen
Durchführung	
Validierung	Schulz, Jansen & Schlotz (2005) – Evaluation:

	<p>Zuverlässigkeit:</p> <p>Cronbachs Alpha: $\alpha = .71 - .91$</p> <p>Retest: $r_{tt} = .63 - .84$</p> <p>Testgültigkeit:</p> <p>Befunde zur faktoriellen Validität liegen vor:</p> <p>Zusammenhänge zu anderen konstrukt-nahen Persönlichkeitsmerkmalen, Beschwerden, Schlafverhalten, chronischen Erkrankungen und Cortisolreaktionen in einer standardisierten Stresssituation</p> <p>Ergebnisse zum Einfluss genetischer Faktoren sowie Geschlechtseffekte</p> <p>Normtabellen sind verfügbar, aber: Stichprobe nicht repräsentativ; es gibt keinen Cut-OFF Wert</p>
Bewertung	<p>Schulz, Jansen & Schlotz (2005):</p> <p>Trennschärfe- und Reliabilitätskoeffizienten erfüllen die teststatistischen Anforderungen</p> <p>Ergebnisse sprechen für die Validität der SRS</p> <p>Rüesch et al., 2006:</p> <p>Weitere Untersuchungen und entsprechende Normierung notwendig</p> <p>Anwendung: klinische Forschung, Erfassung von individuellen Stressreaktivitäten, Evaluation von Anti-Stress-Trainings</p>
Verfügbarkeit	<p>Schulz, Jansen & Schlotz (2005)</p>

4.3.3 PSRS

PSRS (The Perceived Stress Reactivity Scale)	
Autoren	Schlotz, Zoccola, Jansen & Schulz (2011)
Inhalt	Erfasst ebenfalls Stressreaktivität <ul style="list-style-type: none"> - mehrdimensional Faktoren <ul style="list-style-type: none"> - Verlängerte/anhaltende Reaktivität - Reaktivität auf Arbeitsüberlastung - Reaktivität auf soziale Konflikte - Reaktivität auf Versagen/Fehler - Reaktion auf soziale Bewertung
Theorie	Konstrukt Stressreaktivität → Disposition einer Person, auf Belastungen mit entsprechender Stressreaktion zu antworten (entsprechend intensiv/lang/schnell)
Merkmale	<ul style="list-style-type: none"> - Deutsch - Selbstrating, 23 Items, 5-stufig
Durchführung	
Validierung	Schlotz, Zoccola, Jansen & Schulz (2011): Evaluationsstudie USA, UK, Germany (N=2040): Zuverlässigkeit: Cronbachs Alpha: $\alpha = .70 - .80$ Validität: Faktorenstruktur hat sich bestätigt (Erwartete) Zusammenhänge mit: Selbstwirksamkeit, Neurotizismus, chronischem Stress und wahrgenommenem Stress, depressive Symptome, Schlaf
Bewertung	Schlotz, Zoccola, Jansen & Schulz (2011): Ökonomisch und nützliches, einfach zu handhabenes Instrument zur Erfassung der wahrgenommenen Stressreaktivität
Verfügbarkeit	In Schlotz, Zoccola, Jansen & Schulz (2011) (Items in Studie verfügbar)

Appendix

The 23-Item Version of the Perceived Stress Reactivity Scale (PSRS)

Instructions: This questionnaire asks about your reactions to situations which you may have experienced in the past. Three answers are suggested. Please indicate the answer that most closely describes your own reaction in general. Please don't skip any item, even if it may be hard to find the best answer.

PSRS Scale Scoring

01	When tasks and duties build up to the extent that they are hard to manage . . .	<input type="radio"/> I am generally untroubled <input type="radio"/> I usually feel a little uneasy <input type="radio"/> I normally get quite nervous
02	When I want to relax after a hard day at work . . .	<input type="radio"/> This is usually quite difficult for me <input type="radio"/> I usually succeed <input type="radio"/> I generally have no problem at all
03	When I have conflicts with others that may not be immediately resolved . . .	<input type="radio"/> I generally shrug it off <input type="radio"/> It usually affects me a little <input type="radio"/> It usually affects me a lot
04	When I make a mistake . . .	<input type="radio"/> In general, I remain confident <input type="radio"/> I sometimes feel unsure about my abilities <input type="radio"/> I often have doubts about my abilities
05	When I'm wrongly criticized by others . . .	<input type="radio"/> I am normally annoyed for a long time <input type="radio"/> I am annoyed for just a short time <input type="radio"/> In general, I am hardly annoyed at all
06	When I argue with other people . . .	<input type="radio"/> I usually calm down quickly <input type="radio"/> I usually stay upset for some time <input type="radio"/> It usually takes me a long time until I calm down
07	When I have little time for a job to be done . . .	<input type="radio"/> I usually stay calm <input type="radio"/> I usually feel uneasy <input type="radio"/> I usually get quite agitated
08	When I make a mistake . . .	<input type="radio"/> I am normally annoyed for a long time <input type="radio"/> I am normally annoyed for a while <input type="radio"/> I generally get over it easily
09	When I am unsure what to do or say in a social situation . . .	<input type="radio"/> I generally stay cool <input type="radio"/> I often feel warm <input type="radio"/> I often begin to sweat
10	When I have spare time after working hard . . .	<input type="radio"/> It often is difficult for me to unwind and relax <input type="radio"/> I usually need some time to unwind properly <input type="radio"/> I am usually able to unwind effectively and forget about the problems of the day
11	When I am criticized by others . . .	<input type="radio"/> important arguments usually come to my mind when it is too late to still make my point <input type="radio"/> I often have difficulty finding a good reply <input type="radio"/> I usually think of a reply to defend myself
12	When something does not go the way I expected . . .	<input type="radio"/> I usually stay calm <input type="radio"/> I often get uneasy <input type="radio"/> I usually get very agitated
13	When I do not attain a goal . . .	<input type="radio"/> I usually remain annoyed for a long time <input type="radio"/> I am usually disappointed, but recover soon <input type="radio"/> In general, I am hardly concerned at all
14	When others criticize me . . .	<input type="radio"/> I generally don't lose confidence at all <input type="radio"/> I generally lose a little confidence <input type="radio"/> I generally feel very unconfident

Appendix (continued)

15	When I fail at something . . .	<input type="radio"/> I usually find it hard to accept <input type="radio"/> I usually accept it to some degree <input type="radio"/> In general, I hardly think about it
16	When there are too many demands on me at the same time . . .	<input type="radio"/> I generally stay calm and do one thing after the other <input type="radio"/> I usually get uneasy <input type="radio"/> Usually, even minor interruptions irritate me
17	When others say something incorrect about me . . .	<input type="radio"/> I usually get quite upset <input type="radio"/> I normally get a little bit upset <input type="radio"/> In general, I shrug it off
18	When I fail at a task . . .	<input type="radio"/> I usually feel very uncomfortable <input type="radio"/> I usually feel somewhat uncomfortable <input type="radio"/> In general, I don't mind
19	When I argue with others . . .	<input type="radio"/> I usually get very upset <input type="radio"/> I usually get a little bit upset <input type="radio"/> I usually don't get upset
20	When I am under stress . . .	<input type="radio"/> I usually can't enjoy my leisure time at all <input type="radio"/> I usually have difficulty enjoying my leisure time <input type="radio"/> I usually enjoy my leisure time
21	When tasks and duties accumulate to the extent that they are hard to cope with . . .	<input type="radio"/> My sleep is unaffected <input type="radio"/> My sleep is slightly disturbed <input type="radio"/> My sleep is very disturbed
22	When I have to speak in front of other people	<input type="radio"/> I often get very nervous <input type="radio"/> I often get somewhat nervous <input type="radio"/> In general, I stay calm
23	When I have many tasks and duties to fulfill . . .	<input type="radio"/> In general, I stay calm <input type="radio"/> I usually get impatient <input type="radio"/> I often get irritable

Note. The first answer category of each item is coded 0, the second 1, and the third 2. Items marked with "R" are to be reversed (reverse score = 2 - original score). Prolonged Reactivity (PR): 2R, 10R, 20R, 21; Reactivity to Work Overload (RWO): 1, 7, 12, 16, 23; Reactivity to Social Conflict (RSC): 3, 5R, 6, 17R, 19R; Reactivity to Failure (RFa): 8R, 13R, 15R, 18R; Reactivity to Social Evaluation (RSE): 4, 9, 11R, 14, 22R; Perceived Stress Reactivity total score (PSRS-tot): sum of the five scale scores.

Abbildung 32
Items PSRS, Schlotz et al. (2011)

4.4 Physiologische Indikatoren psychischer Bewertung

4.4.1 Kardiovaskuläre Aktivität

Kardiovaskuläre Aktivität (Puls und HRV)	
Variablen	<ul style="list-style-type: none"> - Herzschlagrate (Puls) - Variabilität des Herzschlags (HRV)
Messung	<ul style="list-style-type: none"> - Die Messung erfolgt über Elektroden an Brustkorb, Armen und Beinen - Messung auch über Smartphone (Kamera als Sensor), Smartwatch - Jeder Herzschlag und die Zeit zwischen den Schlägen wird gemessen (für HRV)
Aussage	<ul style="list-style-type: none"> - Valide, objektive und akkurate Messung des körperlichen Stresszustands, sowie der Erholungsfähigkeit

4.4.2 Hautleitfähigkeits-niveau und -reaktion (SCL, SCR)

Hautleitfähigkeits-niveau und -reaktion (SCL, SCR)	
Variablen	<ul style="list-style-type: none"> - SCL (Skin Conductive Level) = Hautleitfähigkeitsniveau - SCR (Skin Conductive Response) = Hautleitfähigkeitsreaktion
Messung	<ul style="list-style-type: none"> - Die Messung erfolgt über zwei Klettelektroden, welche je um das oberste oder mittlere Glied von Zeige- und Mittelfinger einer Hand gewickelt werden.
Aussage	<ul style="list-style-type: none"> - Ebenfalls: Bietet die Möglichkeit die körperliche Stressreaktion objektiv zu messen - Einsetzbar als Bio-Feedback

4.4.3 Cortisolspiegel

Cortisolspiegel	
Variablen	<ul style="list-style-type: none"> - Stresshormon: Cortisol - Wird in Stresssituationen ausgeschüttet → schützt den Körper, denn es sorgt für eine körperliche Anpassung des Körpers an die Anforderung der Situation - Bei dauerhaft hohem Cortisol-Spiegel → negative Folgen, wie z.B. ein geschwächtes Immunsystem
Messung	<ul style="list-style-type: none"> - Messung im Speichel (unkomplizierter) - Messung im Blut
Aussage	<ul style="list-style-type: none"> - Insbesondere chronische Stressbelastung kann sichtbar gemacht werden (konstant hoher Cortisolspiegel)

4.5 Personale Ressourcen

Selbstwirksamkeit

4.5.1 SWE

SWE (Skala zur Allgemeinen Selbstwirksamkeitserwartung)	
Autoren	Jerusalem & Schwarzer (1999)
Inhalt	Erfassung der allgemeinen Selbstwirksamkeit bzw. optimistische Selbstüberzeugungen: Das Vertrauen darauf, eine schwierige Situation auf Basis eigener Kompetenz zu meistern <ul style="list-style-type: none"> - eindimensional
Theorie	Selbstwirksamkeitskonzept Bandura (1977)
Merkmale	<ul style="list-style-type: none"> - Deutsch - Selbstrating, 10 Items, 4-stufig - Ab 12 Jahre
Durchführung	Dauer: 4 min
Validierung	<p>Jerusalem & Schwarzer (1999):</p> <p>Zuverlässigkeit: Cronbachs Alpha $\alpha = .76$ und $.90$</p> <p>Testgültigkeit: Kriterienbezogene Validität erfüllt: z.B. positive Zusammenhänge zu dispositionalem Optimismus, Arbeitszufriedenheit negative Zusammenhänge zu Ängstlichkeit, Depressivität, Burnout, Stresseinschätzungen</p> <p>Normen (T-Werte) sind im Internet zu finden: http://userpage.fu-berlin.de/~health/swe_norm.pdf → (Mittelwert: 29, SD: 4)</p>
Bewertung	Ökonomisches und valides Instrument zur Erfassung der allgemeinen Selbstwirksamkeit
Verfügbarkeit	Frei verfügbar: PT_9001003_SWE_Autorenbeschreibung.pdf

4.5.2 GSE-3

GSE-3 (General Self-Efficacy Short Scale-3)	
Autoren	Doll, Nießen, Schmidt, Rammstedt & Lechner (2021)
Inhalt	<ul style="list-style-type: none"> - Messung der generellen Selbstwirksamkeitsüberzeugung - Eindimensional
Theorie	Selbstwirksamkeitskonzept Bandura (1977)
Merkmale	<ul style="list-style-type: none"> - Deutsch, 3 Items, 5-stufig
Durchführung	Dauer: unter 1 min.
Validierung	<p>Doll et al. (2021): Evaluation UK (N=468); Germany (N=474)</p> <p>Zuverlässigkeit:</p> <p>Cronbachs Alpha $\alpha = .91$;</p> <p>Retest: $r_{tt} = .71$</p> <p>Validität:</p> <p>Konstruktvalidität (Faktorenanalysen) und Kriteriumsvalidität entsprechen Anforderungen</p>
Bewertung	<p>sehr ökonomisch (nur 3 Items)</p> <p>Konstrukt wird dennoch valide abgebildet</p>
Verfügbarkeit	Frei verfügbar (Leibniz-Institut für Sozialwissenschaften)

Items

Table 1

Items of the General Self-Efficacy Short Scale-3 (GSE-3)

No.	Item
1	I can rely on my own abilities in difficult situations.
2	I am able to solve most problems on my own.
3	I can usually solve even challenging and complex tasks well.

Abbildung 33*Items GSE-3, Doll et al. (2021)*

4.5.3 SSA

SSA (Selbstwirksamkeit der sportlichen Aktivität)	
Autoren	Fuchs & Schwarzer (1994)
Inhalt	<ul style="list-style-type: none"> - Erfassung der Selbstwirksamkeit der sportlichen Aktivität: sportliche Aktivität betreiben, auch unter widrigen inneren und externen Umständen - eindimensional
Theorie	Selbstwirksamkeitskonzept Bandura (1977) → Erklärung des sportlichen Verhaltens
Merkmale	<ul style="list-style-type: none"> - deutsch - Selbstrating, 12 Items, 7-stufig
Durchführung	
Validierung	<p>Fuchs & Schwarzer (1994): Evaluation (N=1336)</p> <p>Zuverlässigkeit: Cronbachs Alpha $\alpha = .89$</p> <p>Testgültigkeit: Korrelation mit:</p> <ul style="list-style-type: none"> - generalisierte Selbstwirksamkeit - spezifischen Selbstwirksamkeitserwartungen zur Krebsvorsorge und Ernährung - hohe Korrelation der SSA-Skala mit der Intention zum Sporttreiben ($r = .47$) <p>negative Zusammenhänge:</p> <ul style="list-style-type: none"> - mit Indikatoren des psychischen Befindens (Depressivität, Ängstlichkeit, Ärger) - körperliche Gesundheit (Beschwerdedruck). <p>Trennschärfe → sportliche und unsportliche Personen lassen sich sehr gut trennen</p>
Bewertung	<p>Rüesch et al., 2006:</p> <p>Skala reliables, valides und ökonomisches Verfahren</p> <p>Für Forschungszwecke konzipiert vor allem in den Bereichen Gesundheits- und Sportpsychologie</p>

	Praktische Relevanz (z.B. als diagnostisches Instrument im Bereich der Rehabilitation)
Verfügbarkeit	Frei verfügbar: https://www.sport.uni-freiburg.de/de/institut/Arbeitsbereiche/psychologie/messinstrumente/selbstwirksamkeit-zur-sportlichen-aktivitaet.pdf

Selbstwirksamkeit zur sportlichen Aktivität: SSA-Skala

(Fuchs & Schwarzer, 1994)

Ich bin mir **sicher**, eine geplante Sportaktivität auch dann noch ausüben zu können, wenn...

	gar nicht sicher		viel- leicht		gar sicher								
... ich müde bin.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ich mich niedergeschlagen fühle.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ich Sorgen habe.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ich mich über etwas ärgere.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ich mich angespannt fühle.	1	-	2	-	3	-	4	-	5	-	6	-	7
... Freunde zu Besuch da sind.	1	-	2	-	3	-	4	-	5	-	6	-	7
... andere Personen mit mir etwas unternehmen wollen.	1	-	2	-	3	-	4	-	5	-	6	-	7
... meine Familie/ mein Partner mich beansprucht.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ich niemanden finde, der mit mir Sport treibt.	1	-	2	-	3	-	4	-	5	-	6	-	7
... schlechtes Wetter ist.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ich noch viel Arbeit zu erledigen habe.	1	-	2	-	3	-	4	-	5	-	6	-	7
... ein interessantes Fernsehprogramm läuft.	1	-	2	-	3	-	4	-	5	-	6	-	7

Abbildung 34

Items SSA, Fuchs & Schwarzer (1994)

Kontrollüberzeugung

4.5.4 Pearlin Mastery Scale

Pearlin Mastery Scale	
Autoren	Pearlin & Schooler (1978)
Inhalt	<ul style="list-style-type: none"> - Ausmaß, indem eine Person das eigene Leben unter persönlicher Kontrolle betrachtet - eindimensional
Theorie	
Merkmale	<ul style="list-style-type: none"> - Englisch (Ursprung Spanisch) - Selbstrating, 7 Items, 4-stufig - Summenwerte werden gebildet
Durchführung	
Validierung	<p>Gordon et al. (2018)</p> <ul style="list-style-type: none"> - In spanischer Validierungsstudie konnte sich die Eindimensionalität bestätigen; aber Items luden unterschiedlich stark auf dem Faktor - Ansonsten keine Normierung; keine Validierung
Bewertung	<ul style="list-style-type: none"> - Validierung unzureichend - Keine deutsche Übersetzung
Verfügbarkeit	<ul style="list-style-type: none"> - Frei verfügbar: Pearlin Documentation with IRT.pdf (nlsinfo.org)

Table 1. Pearlin Mastery Scale Items

1. There is really no way I can solve some of the problems I have. (RC)
2. Sometimes I feel that I'm being pushed around in life. (RC)
3. I have little control over the things that happen to me. (RC)
- *4. I can do just about anything I really set my mind to.
5. I often feel helpless in dealing with the problems of life. (RC)
6. What happens to me in the future mostly depends on me.
7. There is little I can do to change many of the important things in my life. (RC)

Note. RC represents item values that require reverse coding prior to scoring. *Item 4 was removed from the scale prior to conducting IRT analyses, due to redundancy with Item 6.

Abbildung 35

Items der Pearlin Mastery Scale (Pearlin & Schooler, 1978)

4.5.5 IPC

IPC	
Autoren	Krampen (1981)
Inhalt	<ul style="list-style-type: none"> - Erfassung von Kontrollüberzeugungen (locus of control) - Mehrdimensional 3 Skalen <ul style="list-style-type: none"> - 1) Internalität - (2) Externalität (Machtlosigkeit) - (3) Externalität (Fatalismus)
Theorie	
Merkmale	- Deutsch, 6-stufig
Durchführung	Keine Informationen
Validierung	IPC scheint validiert, keine Informationen auffindbar
Bewertung	
Verfügbarkeit	Keine Informationen

4.5.6 IE-4-Skala

IE-4-Skala	
Autoren	Kovaleva, Beierlein, Kemper, Rammstedt (2014)
Inhalt	Erfassung der internale und externale Kontrollüberzeugung
Theorie	Kontrollüberzeugung nach Rotter (1966)
Merkmale	Deutsch Selbstrating, 4 Items, 5-stufig
Durchführung	Dauer: unter 1 min.
Validierung	Kovaleva et al. (2014): N=1134 Zuverlässigkeit: Reliabilität: Retest-Reliabilität = .56 bis .64 Cronbachs Alpha = .53 bis .71 Testgültigkeit: Hinweise auf die inhaltliche Validität, faktorielle Validität, konvergente, divergente und prädiktive Validität
Bewertung	Sozialwissenschaftliche Untersuchungen unterschiedlichster Art
Verfügbarkeit	Frei verfügbar: ZIS (gesis.org)

Items

Subskalen: IKÜ = internale Kontrollüberzeugung, EKÜ = externale Kontrollüberzeugung

Nr.	Items (deutsch)	Items (englisch)	Subskalen
1	Ich habe mein Leben selbst in der Hand.	I'm my own boss.	IKÜ
2	Wenn ich mich anstrengte, werde ich auch Erfolg haben.	If I work hard, I will succeed.	IKÜ
3	Egal ob privat oder im Beruf: Mein Leben wird zum großen Teil von anderen bestimmt.	Whether at work or in my private life: What I do is mainly determined by others.	EKÜ
4	Meine Pläne werden oft vom Schicksal durchkreuzt.	Fate often gets in the way of my plans.	EKÜ

Antwortvorgaben

Die vier Items werden mittels einer Skala mit fünfstufigem Antwortformat von "trifft gar nicht zu" (1) bis "trifft voll und ganz zu" (5) beantwortet.

Abbildung 36

Items der IE-4-Skala (Kovaleva et al., 2014)

4.5.7 ASF-E

ASF-E	
Autoren	Poppe, Pelster & Pelster (2005)
Inhalt	<ul style="list-style-type: none"> - Erfassung des Attributionsstils: Welches Erklärungsmuster wird für positive und negative Ereignisse bevorzugt? - Attributionsstil als stabile Persönlichkeitseigenschaft
Theorie	Attributionsstil
Merkmale	<ul style="list-style-type: none"> - Deutsch - Selbstrating, 16 Items (8 pos., 8 neg.)
Durchführung	Dauer: 20-40 min
Validierung	<p>Poppe, Pelster & Pelster (2005)</p> <p>Zuverlässigkeit: $\alpha = .88-.89$</p> <p>Testgültigkeit: Konstruktvalidität:</p> <ul style="list-style-type: none"> - Kennwerte des ASF-E korrelieren mit Depressionskennwerten, insbesondere bei Personen, die zuvor negative Lebensereignisse erlebt haben. - erwartungsgemäße Veränderung nach Psychotherapie <p>Normen vorhanden (ca. N=1000)</p>
Bewertung	<ul style="list-style-type: none"> - Abklärung der Ursachen einer Depression und von Selbstwert-, Motivations- und Leistungsdefiziten - Evaluation kognitiv-verhaltenstherapeutischer Interventionen. - Personalauswahl
Verfügbarkeit	Poppe, Pelster & Pelster, 2005 (→ Hogrefe)

4.5.8 SCS

SCS (Self-Control Schedule)	
Autoren	Seipel (2002)
Inhalt	<ul style="list-style-type: none"> - Erfassung von Selbstkontrolle - Mehrdimensional Subskalen: <ul style="list-style-type: none"> - Impulsivity (I) - Simple Tasks (ST) - Risk-Seeking (RS) - Physical Activities (PA) - Self-Centered (SC) - Temper (T)
Theorie	Theory of Crime (GTOC; Gottfredson und Hirschi 1990), Kontrolltheorien sechs Bereiche von low Self-Control: <ol style="list-style-type: none"> 1. eine starke Hier- und Jetzt-Orientierung (Impulsivity), 2. eine geringe Sorgfalt, Persistenz und Verlässlichkeit (Simple Tasks), 3. eine starke Abenteuerlust (Risk-Seeking), 4. ein starkes Interesse an körperlicher Aktivität (Physical Activities), 5. eine starke Tendenz, selbstbezogen, indifferent gegenüber anderen und unsensibel zu sein (Self-Centered) 6. eine geringe Frustrationstoleranz (Temper). <p>Es wird angenommen, dass Personen mit einer geringen Self-Control dazu neigen, eine Vielzahl verschiedener Formen abweichenden Verhaltens auszuführen, sofern sich entsprechende Gelegenheiten dazu bieten.</p>
Merkmale	<ul style="list-style-type: none"> - Deutsch - Selbstrating, 23 Items, 5-stufig
Durchführung	
Validierung	Seipel (2002): Zuverlässigkeit: Cronbachs Alpha = .44 bis .76 Testgültigkeit: Validitätskriterien erfüllt (Faktoren in konfirmatorischer Faktorenanalyse bestätigt)

Bewertung	Fragebogen ist entsprechend validiert; aber keine Normierung
Verfügbarkeit	Frei verfügbar: ZIS (gesis.org)

Instruktion

Nun einige Fragen zu Ihrer Person: Inwieweit treffen die folgenden Aussagen auf Sie zu? Bitte kreisen Sie in jeder Zeile die für Sie zutreffende Antwortziffer ein!

Items			Aus-	Sub-
Nr.	Item		schluss*	skala
A	Ich handle oft spontan, denke aber trotzdem mit.			I
B	Ich mache mir nicht viele Gedanken über die Zukunft.			I
C	Ich lebe jetzt und hier und tue was immer mir Spaß bringt, auch auf Kosten eines entfernteren Ziels.			I
K	Ich beschäftige mich mehr mit dem, was mir in naher Zukunft passiert, als mit Dingen, die noch weiter entfernt sind.	*		I
E	Ich vermeide oft Aufgaben, von denen ich weiß, dass sie schwierig werden.			ST
F	Wenn Dinge schwierig werden, neige ich dazu, aufzugeben oder mich zurückzuziehen.			ST
G	Die Dinge im Leben, die mir am leichtesten fallen, bringen mir am meisten Spaß.			ST
H	Schwere Aufgaben, die mich voll beanspruchen, mag ich überhaupt nicht.			ST
D	Manchmal riskiere ich etwas, nur um Spaß zu haben.			RS
J	Hin und wieder setze ich mich kleinen Risiken aus, um mich auszutesten.			RS
L	Ich finde es manchmal aufregend, Sachen zu machen, für die ich Ärger bekommen könnte.			RS
M	Aufregung und Abenteuer sind für mich wichtiger als Sicherheit.			RS
O	Wenn ich die Wahl hätte, würde ich meistens körperliche Aktivitäten geistigen vorziehen.			PA
Q	Ich gehe lieber aus, als dass ich ein Buch lese.			PA
R	Ich glaube, dass ich mehr Action brauche als andere meines Alters.			PA
Y	Ich fühle mich meistens besser, wenn ich körperlich aktiv bin, als wenn ich dasitze und denke.	*		PA
S	Ich denke zuerst an mich, ohne viel Rücksicht auf andere zu nehmen.			SC
T	Ich bin Personen gegenüber, die Probleme haben, nicht gerade aufgeschlossen.			SC
V	Wenn mein Verhalten andere Menschen stört, dann ist das ihr Problem.			SC
P	Wenn ich so richtig wütend bin, sollte man mich lieber in Ruhe lassen.			T
W	Ich verliere ziemlich leicht die Beherrschung.			T
X	Wenn ich wütend auf jemand bin, dann beleidige ich ihn eher, als darüber zu sprechen, was mich wütend macht.			T
Z	Wenn ich eine ernsthafte Auseinandersetzung mit jemandem habe, ist es normalerweise schwierig für mich, ruhig zu reden und nicht zu explodieren.			T

Abbildung 37

Items des SCS (Seipel, 2002)

Optimismus

4.5.9 H-Skalen

Hoffnungslosigkeit-Skalen	
Autoren	Krampen (1979)
Inhalt	Die H-Skalen erfassen wesentliche Aspekte der kognitiven Triade der Depression mit den negativen Erwartungen einer Person <ul style="list-style-type: none"> - a) über sich selbst - b) über die Umwelt und - c) über ihr zukünftiges Leben
Theorie	
Merkmale	<ul style="list-style-type: none"> - deutsche Übersetzung - Selbstrating, 20 Items, ja-nein Kategorien (richtig/falsch)
Durchführung	
Validierung	Krampen (1979): Evaluationsstudie (N=221) Zuverlässigkeit: Cronbachs Alpha $\alpha = .87$ Testhalbierungsreliabilität: rtt = .87 Testgültigkeit: Validitätsprüfungen liegen vor: Korrelation mit externen Kontrollüberzeugungen und Machiavellismus
Bewertung	Insbesondere Klinischer Kontext
Verfügbarkeit	Frei verfügbar in Krampen (1979): 1979 Hoffnungslosigkeit bei stationären Patienten.pdf (uni-trier.de)

Tab. 1: Polungen (P), Schwierigkeitsindizes (p_i) und Trennschärfekoeffizienten (r_{it-i}) der Items der H-Skala (N = 221)

	p^a	P_i	r_{it-i}
1. Ich blicke mit Optimismus und Begeisterung in die Zukunft.	R	.36	.59
2. Häufig möchte ich alles hinschmeißen, weil ich es doch nicht besser machen kann.	F	.62	-.55
3. Wenn einmal alles schief läuft, geht es mir besser, wenn ich daran denke, daß es ja auch wieder aufwärts gehen wird.	R	.17	.32
4. Ich kann mir nicht vorstellen, wie mein Leben in 10 Jahren aussehen wird.	F	.34	-.52
5. Ich habe genug Zeit, um die Sachen, die mir Spaß machen, zu tun.	R	.45	.27
6. Ich glaube, daß ich in der Zukunft mit dem, was mich im Moment am meisten beschäftigt, Erfolg haben werde.	R	.28	.55
7. Die Zukunft liegt für mich im Dunkel.	F	.56	-.69
8. Ich erwarte, in meinem Leben mehr Schönes zu erleben als der durchschnittliche Mensch.	R	.50	.37
9. Ich kriege einfach keine richtigen Chancen im Leben.	F	.70	-.59
10. Meine Erfahrungen sind eine gute Vorbereitung für künftige Probleme.	R	.15	.42
11. Alles, was ich so im Moment vor mir liegen sehe, ist eher unschön als schön und angenehm.	F	.52	-.59
12. Ich glaube nicht, daß ich jemals das im Leben bekomme, was ich mir in Wahrheit wünsche.	F	.49	-.67
13. Ich glaube, daß ich in künftigen Zeiten glücklicher sein werde als heute.	R	.47	.27
14. Die Dinge laufen einfach nicht so, wie ich es gerne hätte.	F	.38	-.69
15. Ich setze große Hoffnungen in die Zukunft.	R	.33	.52
16. Ich bekomme einfach nie das, was ich will; es ist also Unsinn, überhaupt noch etwas zu wollen	F	.83	-.60
17. Es ist sehr unwahrscheinlich, daß das Leben mir noch Befriedigung und Freude bringt.	F	.80	-.49
18. Die Zukunft erscheint mir sehr unsicher.	F	.52	-.75
19. Das Leben wird mir noch viel mehr schöne Zeiten bringen als schlechte.	R	.33	.58
20. Es nützt nichts, etwas anzustreben, das ich gerne hätte, da ich es wahrscheinlich ja doch nicht erreiche.	F	.70	-.53

a Bei der Polung steht F für "falsch" und R für "richtig" bei der Beantwortung der Items; die Schwierigkeitsindizes beziehen sich auf die Antwortkategorie "falsch".

Abbildung 38

Items der H-Skala (Krampen, 1979)

Resilienz

4.5.10 RS

RS (Resilienzskala)	
Autoren	Wagnild und Young (1993) Deutsch: Schumacher, Leppert, Gunzelmann, Strauß & Brähler (2005)
Inhalt	<ul style="list-style-type: none"> - Selbstbeurteilungsverfahren zur Erfassung der Bewältigungsfähigkeit (persönliche Kompetenz und individuelle Widerstandsfähigkeit) - eindimensional
Theorie	Resilienz als dynamische Fähigkeit pos. Affekte situations- und kontextadäquat kontrollieren und modifizieren zu können
Merkmale	<ul style="list-style-type: none"> - Deutsch - Selbstrating, 25/13/11 Items, 7-stufig
Durchführung	Dauer: 10 min
Validierung	Validierungsstudie (Schumacher et al., 2005): Evaluation (N=2031) Reliabilität und Validität entsprechen den teststatistischen Anforderungen Zuverlässigkeit: Cronbachs Alpha: $\alpha = .82 - .95$ Testgültigkeit: Zusammenhänge z.B. mit Allgemeiner Selbstwirksamkeit, Prädiktor subjektiver Körperbeschwerden Normwerte sind vorhanden
Bewertung	Unklarheiten bezüglich des Konzepts Resilienz (Rüesch et al., 2006)
Verfügbarkeit	RS-11 in Schumacher et al., 2005: Resilienzskala ZKPPP rev.doc (bgm-ost-schweiz.ch)

Resilienzskala RS-11

Im folgenden Fragebogen finden Sie eine Reihe von Feststellungen. Bitte lesen Sie sich jede Feststellung durch und kreuzen Sie an, wie sehr die Aussagen im Allgemeinen auf Sie zutreffen, d.h. wie sehr Ihr übliches Denken und Handeln durch diese Aussagen beschrieben wird.

		1 = nein					7 = ja	
		Ich stimme nicht zu stimme völlig zu						
		1	2	3	4	5	6	7
1	Wenn ich Pläne habe, verfolge ich sie auch.	1	2	3	4	5	6	7
2	Normalerweise schaffe ich alles irgendwie.	1	2	3	4	5	6	7
3	Es ist mir wichtig, an vielen Dingen interessiert zu bleiben.	1	2	3	4	5	6	7
4	Ich mag mich.	1	2	3	4	5	6	7
5	Ich kann mehrere Dinge gleichzeitig bewältigen.	1	2	3	4	5	6	7
6	Ich bin entschlossen.	1	2	3	4	5	6	7
7	Ich behalte an vielen Dingen Interesse.	1	2	3	4	5	6	7
8	Ich finde öfter etwas, worüber ich lachen kann.	1	2	3	4	5	6	7
9	Normalerweise kann ich eine Situation aus mehreren Perspektiven betrachten.	1	2	3	4	5	6	7
10	Ich kann mich auch überwinden, Dinge zu tun, die ich eigentlich nicht machen will.	1	2	3	4	5	6	7
11	In mir steckt genügend Energie, um alles zu machen, was ich machen muss.	1	2	3	4	5	6	7

Abbildung 39

Items der RS-11 (Schumacher et al., 2005)

Kohärenzerleben

4.5.11 SOC-Skala

SOC-Skala	
Autoren	Antonovsky (1979) SOC-L9 (Schumacher & Brähler, 2004)
Inhalt	<ul style="list-style-type: none"> - Erfassung der Komponenten des Kohärenzgefühls - Personale, dispositionelle Bewältigungsressource - eindimensional
Theorie	Salutogenese (Antonovsky, 1979)
Merkmale	<ul style="list-style-type: none"> - Es gibt verschiedene Versionen - SOC-L9: Selbstrating, 9-29 Items - Ab 18 Jahren
Durchführung	Dauer: 5 min
Validierung	<p>Verschiedene Versionen z.B. SOC-L9 (Schumacher & Brähler, 2004)</p> <p>Zuverlässigkeit: Reliabilität: $\alpha=.87$</p> <p>Testgültigkeit: Die Konstruktvalidität und Eindimensionalität der SOC-L9 wurde faktorenanalytisch gesichert. Die Korrelation der SOC-L9 mit der Langfassung der SOC-Skala mit 29 Items (SOC-29) beträgt $r = .94$. Normstichprobe vorhanden (N=1944)</p>
Bewertung	Validiert, normiert
Verfügbarkeit	Verfügbar in Schumacher & Brähler (2004): iqpr-assessment: Assessmentkatalog -> Detailansicht (assessment-info.de)

4.6 Planungshilfen und Empfehlungen zu Stress-Erfassung

Entsprechend des transaktionalen Stressmodells von Lazarus ist die subjektive Bewertung der Reize entscheidend für die Stressreaktion. Des Weiteren reagiert der Körper unabhängig von der Art des Stressors mit einer körperlichen Anpassungsreaktion (Gesundheitsberichterstattung des Bundes, 2022), sodass objektive Messdaten (Cortisolspiegel, Hautleitfähigkeit, und kardiovaskuläre Aktivität) aussagekräftig sein können, jedoch nicht ausreichen, um den individuellen Beanspruchungsgrad abzubilden. Dementsprechend ergibt sich eine Notwendigkeit einer multimodalen Erfassung verschiedener Stressindikatoren (Kombination aus objektiven Parametern und subjektiver Einschätzung), um eine möglichst adäquate Erfassung des subjektiven Stresslevels zu gewährleisten.

Abbildung 41 zeigt eine Übersicht der Stress-Erhebungsmethoden und versucht deren Klassifikation. Eine Empfehlung kann nicht allgemein gegeben werden, sondern hängt von den apparativen und theoretischen Voraussetzungen des jeweiligen Studienziels ab.

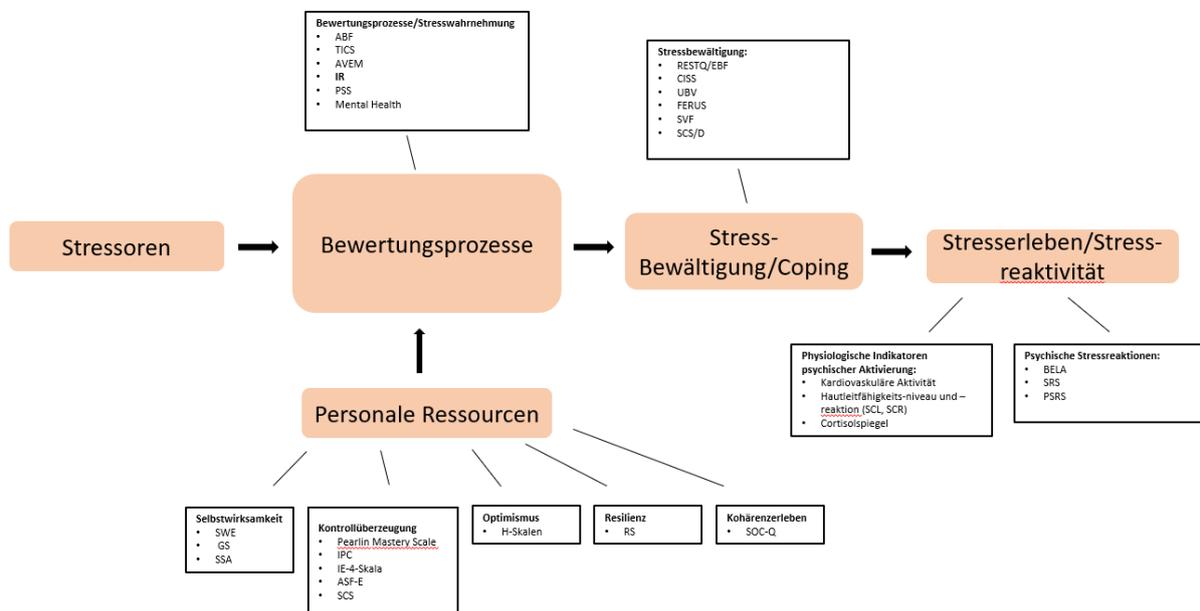


Abbildung 40

Messmethoden auf Basis des transaktionalen Stressmodells (in Anlehnung an Lazarus & Folkman, 1987)

5 Literatur

- Adams, S. (1998). Practical considerations for measuring situational awareness. In *Proceedings for the Third Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment* (S. 157–164).
- Aldrich, T. B. & Szabo, S. M. (1986). A methodology for predicting crew workload in new weapon systems. *Proceedings of the Human Factors Society Annual Meeting*, 30(7), 633–637. <https://doi.org/10.1177/154193128603000705>
- Alimohammadi, I., Azadi, N., Damiri, Z., Ebrahimi, H. & Yeganeh, R. (2020). The effect of age on driving performance in Iran using driving simulator. *Archives of Trauma Research*, 9(3), 116–123. https://doi.org/10.4103/atr.atr_113_19
- Antin, J. F. & Wierwille, W. W. (1984). Instantaneous measures of mental workload: An initial investigation. *Proceedings of the Human Factors Society Annual Meeting*, 1(28), 6–10. <https://doi.org/10.1177/154193128402800104>
- Antonovsky, A. (1979). *Health, stress, and coping: New Perspectives on Mental and Physical Well-Being*, 12-37. San Francisco: Jossey-Bass.
- Badke-Schaub, P., Hofinger, G. & Lauche, K. (2008). *Human Factors*. Springer. <https://doi.org/10.1007/978-3-540-72321-9>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191–215.
- Bartenwerfer, H. (1969). Einige praktische Konsequenzen aus der Aktivierungstheorie. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 16(2), 195–222.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1987). *Beck depression inventory*. New York: Harcourt Brace Jovanovich.
- Boucsein, W. (1973). *Analyse einiger psychologischer Testverfahren zur Erfassung von Persönlichkeitsmerkmalen*. Unpublished Report of the Psychological Institute, Düsseldorf.
- Boumann, H., Cyrol, D., Biella, M. & Sammito, S. (2022). *Übersicht zu physiologischen und psychologischen Parametern zur Belastungsmessung und deren Nutzbarkeit im Setting „air systems“ (inkl. Fragebögen)*. Stand: 02.11.2022. Forschungsbericht 2022/01-m³i-wasp. Köln-Lind.
- Bowers, C. A., Barnett, J., Weaver, J. & Stout, R. J. (1998). Empirical validation of the SALIANT methodology. In *Proceedings of the North Atlantic Treaty Organization Research and Technology Organization Meeting 4*, 12.
- Brantley, Ph., Waggoner, C. D. Jones, G. N. & Rappaport, N. B. (1987). A daily stress inventory: Development, reliability, and validity. *Journal of Behavioral Medicine*, 10, 61-74.
- Brennan, S. D. (1992). An experimental report on rating scale descriptor sets for the instantaneous self assessment (ISA) recorder. *DRA Maritime Command and Control Division. DRA Technical Memorandum (CAD5)*, 92017.
- Brookhuis, K. A., van Driel, C. J. G., Hof, T., van Arem, B. & Hoedemaeker, M. (2009). Driving with a congestion assistant; mental workload and acceptance. *Applied Ergonomics*, 40(6), 1019–1025.
- Bryant-Lees, K. B., Prince, L., Goodman, T., Chappelle, W., & Thompson, B. (2021). Sources of stress and psychological health outcomes for remotely piloted aircraft operators: A comparison across career fields and major commands. *Military Medicine*, 186(7-8), e784-e795.
- Buchwald, P. & Hobfoll, S. E. (2013). Die Theorie der Ressourcenerhaltung. Implikationen für den Zusammenhang von Stress und Kultur. In P. Genkova, T. Ringeisen & F. T. L. Leong (Eds.), *Handbuch Stress und Kultur. Interkulturelle*

- und kulturvergleichende Perspektiven (S. 127–138). Springer Fachmedien Wiesbaden.
- Budescu, D. V., Zwick, R. & Rapoport, A. (1986). A comparison of the eigenvalue method and the geometric mean procedure for ratio scaling. *Applied Psychological Measurement*, 10(1), 69–78.
<https://doi.org/10.1177/014662168601000106>
- Cain, B. (2007). Review of the mental workload literature. *Defence Research And Development*.
- Carretta, T. R., Perry, D. C. & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *The International Journal of Aviation Psychology*, 6(1), 21–41.
https://doi.org/10.1207/s15327108ijap0601_2
- Casali, J. G. & Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors*, 25(6), 623–641.
- Casali, J. G. & Wierwille, W. W. (1984). On the measurement of pilot perceptual workload: a comparison of assessment techniques addressing sensitivity and intrusion issues. *Ergonomics*, 27(10), 1033–1050.
- Casner, S. M. & Gore, B. F. (2010). Measuring and evaluating workload: A primer. *NASA Technical Memorandum*, 216395.
- Charkhandaz Yeganeh, R., Alimohammadi, I., Abolghasemi, J., Damiri, Z., Parsazadeh, B. & Rahmani, N. (2016). Validity and reliability of Verbal Online Subjective Opinion (VOSO) and Modified Cooper-Harper scales in measuring of mental workload. *Journal of Occupational Hygiene Engineering*, 3(3), 24–31.
<https://doi.org/10.21859/johe-03038>
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385–396.
- Cooper, G. E. & Harper Jr, R. P. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities* (No. AGARD-567). Advisory group for aerospace research and development.
- Cummings, M. L., Pina, P. E. & Donmez, B. (2008). *Selecting metrics to evaluate human supervisory control applications*. MIT Humans and Automation Laboratory.
- Dennehy, K. (1997). *Cranfield Situation Awareness Scale: Users' manual* (Bd. 9702). College of Aeronautics Report.
- De Waard, D. (1996). *The measurement of drivers' mental workload* [Doctoral Dissertation, Groningen University].
- Dishman, D. (2019). *Adaptation and validation of the Situation Awareness Global Assessment Technique for student registered nurse anesthetists* [Doctoral Dissertation]. Virginia Commonwealth University.
<https://doi.org/10.25772/1YEM-0281>
- Doll, E. S., Nießen, D., Schmidt, I., Rammstedt, B. & Lechner, C. M. (2021). The General Self-Efficacy Short Scale–3 (GSE-3): An English-language adaptation. In *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)* (Vol. 10).
- Durso, F. T., Bleckley, M. K. & Dattel, A. R. (2006). Does Situation Awareness add to the validity of cognitive tests? *Human Factors*, 48(4), 721–733.
<https://doi.org/10.1518/001872006779166316>
- 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>

- Edgar, G. K., Di Catherwood, Baker, S., Sallis, G., Bertels, M., Edgar, H. E., Nikolla, D., Buckle, S., Goodwin, C. & Whelan, A. (2018). Quantitative Analysis of Situation Awareness (QASA): Modelling and measuring situation awareness using signal detection theory. *Ergonomics*, 61(6), 762–777. <https://doi.org/10.1080/00140139.2017.1420238>
- Edgar, G. K., Smith, A. J., Stone, H. E., Beetham, D. L. & Pritchard, C. (2000). QUASA: QUantifying and Analysing Situational Awareness. In *IMCD People in Digitized Command and Control Symposium*. RMCS Shrivenham, UK. <http://eprints.glos.ac.uk/id/eprint/590>
- Eggemeier, F. T., Wilson, G. F., Kramer, A. F. & Damos, D. L. (1991). Workload assessment in multi-task environments. In D. L. Damos (Hg.), *Multiple-task performance* (S. 207–216). Taylor & Francis.
- Endler, N. S., Parker, J. D. & Summerfeldt, L. J. (1999). Coping with health problems: developing a reliable and valid multidimensional measure. *Psychological Assessment*, 10(3), 195–205.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society Annual Meeting*, 32(2), 97–101. <https://doi.org/10.1177/154193128803200221>
- Endsley, M. R. (1990). Predictive utility of an objective measure of Situation Awareness. *Proceedings of the Human Factors Society Annual Meeting*, 34(1), 41–45. <https://doi.org/10.1177/154193129003400110>
- Endsley, M. R. (2021). A systematic review and meta-analysis of direct objective measures of situation awareness: a comparison of SAGAT and SPAM. *Human Factors*, 63(1), 124–150.
- Endsley, M. R. & Garland, D. J. (Hg.). (2008). *Situation awareness analysis and measurement (Re-printed)*. CRC Press.
- Endsley, M. R., Selcon, S. J., Hardiman, T. D. & Croft, D. G. (1998). A comparative analysis of SAGAT and SART for evaluations of situation awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(1), 82–86. <https://doi.org/10.1177/154193129804200119>
- Endsley, M. R., Sollenberger, R. & Stein, E. (2000). Situation awareness: A comparison of measures. In *Proceedings of the Human Performance, Situation Awareness and Automation: User-centered design for the new millenium* (S. 15–19).
- 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*, 63(1), 124–150. <https://doi.org/10.1177/0018720819875376>
- Erdmann, G., & Janke, W. (2008). *Stressverarbeitungsfragebogen: SVF; Stress, Stressverarbeitung und ihre Erfassung durch ein mehrdimensionales Testsystem*. Hogrefe.
- Farmer, E. W., Belyavin, A. J., Jordan, C. S., Bunting, A. J., Tattersall, A. J. & Jones, D. M. (1995). *Predictive workload assessment: Final report*. Farnborough, Hampshire. Defence Research Agency.
- Farmer, E. W., Jordan, C. S., Belyavin, A. J., Bunting, A. J., Tattersall, A. J. & Jones, D. M. (1995). *Dimensions of operator workload: Final report*. Farnborough, UK. Defence Research Agency, 58.
- Fuchs, R. & Schwarzer, R. (1994). Selbstwirksamkeit zur sportlichen Aktivität: Reliabilität und Validität eines neuen Meßinstruments [Self-efficacy toward physical exercise: Reliability and validity of a new instrument]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15(3), 141–154.
- Funke, G., Knott, B., Mancuso, V. F., Strang, A., Estep, J., Menke, L. & Miller, B. (2013). Evaluation of subjective and EEG-based measures of mental

- workload. In C. Stephanidis (Hg.), *HCI International 2013 – Posters' Extended Abstracts*, (S. 412–416). Springer.
- Gatsoulis, I. (2008). *Performance metrics and human-robot interaction for teleoperated systems* [Doctoral Dissertation]. University of Leeds.
- Gawron, V. J. (2019). *Human performance and situation awareness measures* (3rd edition). CRC Press/Taylor & Francis Group.
- Gawron, V. J., Weingarten, N., Hughes, T. & Adams, S. (1999). Verifying Situational Awareness associated with flight symbology. *37th Aerospace Sciences Meeting and Exhibit*, 40. <https://doi.org/10.2514/6.1999-1093>
- Gesundheitsberichterstattung des Bundes (2022). Stress. https://www.gbe-bund.de/gbe/abrechnung.prc_abr_test_lo-gon?p_uid=gast&p_aid=0&p_knoten=FID&p_sprache=D&p_suchstring=8612 [accessed 08.06.2022]
- Goillau, P. J. & Kelly, C. J. (1997). MAIvern Capacity Estimate (MACE): A proposed cognitive measure for complex systems. In D. Harris (Hg.), *Engineering psychology and cognitive ergonomics: Volume 1: Transportation systems* (S. 219–225). Ashgate Publishers.
- Gordon, J. R., Malcarne, V. L., Roesch, S. C., Roetzheim, R. G. & Wells, K. J. (2018). Structural validity and measurement invariance of the Pearlin Mastery Scale in Spanish-speaking primary care patients. *Evaluation and the Health Professions*, 41(3), 393–399.
- Gottfredson, M. R. & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.
- Grosjean, V. & Terrier, P. (1999). Temporal awareness: Pivotal in performance? *Ergonomics*, 42(11), 1443–1456. <https://doi.org/10.1080/001401399184802>
- Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), 904–908. <https://doi.org/10.1177/154193120605000909>
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*, (Bd. 52, S. 139–183). [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- 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)
- Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner Jr, A. C., Zaklade, A. L. & Christ, R. E. (1992). Comparison of four subjective workload rating scales. *Human Factors*, 34(4), 429–439.
- Hobfoll, S. E., & Buchwald, P. (2004). Die Theorie der Ressourcenerhaltung und das multiaxiale Copingmodell – eine innovative Stresstheorie. In P. Buchwald, C. Schwarzer & S. E. Hobfoll (Eds.), *Stress gemeinsam bewältigen – Ressourcenmanagement und multi-axiales Coping* (S. 11–26). Göttingen: Hogrefe.
- Huey, B. & Wickens, C. D. (1993). *Workload transition: Implications for Individual and team performance*. National Academy Press, 54–93.
- Hupfeld, J. & Ruffieux, N. (2011). Validierung einer deutschen Version der Self-Compassion Scale (SCS-D) [Validation of a German version of the Self-Compassion Scale (SCS-D)]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 40(2), 115–23.
- Jack, M. (2007). *Fragebogen zur Erfassung von Ressourcen und Selbstmanagementfähigkeiten (FERUS); Manual*. Hogrefe.
- Jain, S., Shapiro, S. L., Swanick, S., Roesch, S. C., Mills, P. J., Bell, I., & Schwartz, G. E. (2007). A randomized controlled trial of mindfulness meditation versus

- relaxation training: effects on distress, positive states of mind, rumination, and distraction. *Annals of Behavioral Medicine*, 33(1), 11-21.
- Jeannot, E., Kelly, C. J. & Thompson, D. (2003). *The development of situation awareness measures in ATM systems*. EUROCONTROL, the European Organisation for the Safety of Air Navigation.
- Jennings, S., Cheung, B., Rupert, A. H., Schultz, K. & Craig, G. (2004). Flight-test of a tactile situational awareness system in a land-based deck landing task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(1), 142–146. <https://doi.org/10.1177/154193120404800131>
- Jerusalem, M., & Schwarzer, R. (1999). Allgemeine Selbstwirksamkeitserwartung. *Skalen zur Erfassung von Lehrer- und Schülermerkmalen*, 13-14.
- Jordan, C. S. & Brennan, S. D. (1992). *Instantaneous self-assessment of workload technique (ISA)*. Defence Research Agency, Portsmouth. (kein Zugriff).
- Kälin, W. & Semmer, N. (2020). *CISS: Coping-Inventar zum Umgang mit Stress-Situationen. Deutschsprachige Adaptation des Coping Inventory for Stressful Situations (CISS [TM]) von Norman S. Endler und James D. A. Parker*. Bern: Hogrefe
- Kallus, K. W. & Kellmann, M. (Eds.). (2016). *The recovery-stress questionnaires: User manual* (S. 360). Frankfurt/Main: Pearson.
- Kanfer, F. H., Reinecker, H. & Schmelzer, D. (1969). *Selbstmanagementtherapie* (2., überarbeitete Aufl.). Berlin: Springer.
- Kintz, J. R., Banerjee, N. T., Zhang, J. Y., Anderson, A. P., & Clark, T. K. (2023). Estimation of Subjectively Reported Trust, Mental Workload, and Situation Awareness Using Unobtrusive Measures. *Human Factors*, 65(6), 1142–1160. <https://doi.org/10.1177/00187208221129371>
- Kirwan, B., Evans, A., Donohoe, L., Kilner, A., Lamoureux, T., Atkinson, T. & MacKendrick, H. (1997). Human factors in the ATM system design life cycle. In *FAA/Eurocontrol ATM R&D Seminar 1997* (S. 16–20).
- Klein, E., Brähler, E., Dreier, M., Reinecke, L., Müller, K., Schmutzer, G., . . . & Beutel, M. (2016). The German version of the Perceived Stress Scale: Psychometric characteristics in a representative German community sample. *BMC Psychiatry*, 16, 159. <https://doi.org/10.1186/s12888-016-0875-9>
- Kovaleva, A., Beierlein, C., Kemper, C. J. & Rammstedt, B. (2014). *Internale-Externale-Kontrollüberzeugung-4 (IE-4). Zusammenstellung sozialwissenschaftlicher Items und Skalen*. GESIS - Leibniz-Institut für Sozialwissenschaften
- Krampen, G. (1979). Hoffnungslosigkeit bei stationären Patienten: Ihre Messung durch einen Kurzfragebogen (H-Skala). *Medizinische Psychologie*, 5(1), 39-49.
- Krampen, G. (1981). Differentialpsychologische Korrelate von Kontrollüberzeugungen [Personality correlates of three aspects of locus of control]. *Diagnostica*, 27(1), 78–80.
- Langan-Fox, J., Sankey, M. J. & Canty, J. M. (2009). Human factors measurement for future air traffic control systems. *Human Factors*, 51(5), 595–637. <https://doi.org/10.1177/0018720809355278>
- Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, 1(3), 141-170.
- Lee, K., Kerns, K., Bone, R. & Nickelson, M. (2001). Development and validation of the controller acceptance rating scale (CARS): Results of empirical research. In *Proceedings of the 4th USA/Europe Air Traffic Management R&D Seminar*.
- Leggatt, A. (2005). Validation of the ISA (Instantaneous Self Assessment) subjective workload tool. In P. D. Bust & P. T. McCabe (Hg.), *Contemporary Ergonomics 2005* (S. 74–78). Taylor & Francis.

- Li, A. W., & Goldsmith, C. A. W. (2012). The effects of yoga on anxiety and stress. *Alternative Medicine Review, 17*(1).
- Litzcke, S. M., & Schuh, H. (2010). Stressentstehung und Stressreaktion. In *Stress, Mobbing und Burnout am Arbeitsplatz* (S. 5-32). Berlin, Heidelberg: Springer.
- Loft, S., Bowden, V., Braithwaite, J., Morrell, D. B., Huf, S. & Durso, F. T. (2015). Situation awareness measures for simulated submarine track management. *Human Factors, 57*(2), 298–310. <https://doi.org/10.1177/0018720814545515>
- Longo, L., Wickens, C. D., Hancock, P. A., & Hancock, G. M. (2022). Human mental work-load: A survey and a novel inclusive definition. *Frontiers in Psychology, 13*. doi: 10.3389/fpsyg.2022.883321
- Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D. & Linton, P. M. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies*. Defense Technical Information Center. <https://doi.org/10.21236/ada212879>
- Mansikka, H., Virtanen, K. & Harris, D. (2019). Comparison of NASA-TLX scale, modified Cooper–Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks. *Ergonomics, 62*(2), 246–254.
- Manzoni, G. M., Pagnini, F., Castelnovo, G., & Molinari, E. (2008). Relaxation training for anxiety: a ten-years systematic review with meta-analysis. *BMC Psychiatry, 8*(1), 1-12.
- Matthews, M. D., Beal, S. A. & Pleban, R. J. (2002). *Situation Awareness in a virtual environment: description of a subjective assessment scale* (Vol. 2). US Army Research Institute for the Behavioral and Social Sciences. <https://doi.org/10.21236/ada399408>
- Matthews, M. D., Eid, J., Johnsen, B. H. & Boe, O. C. (2011). A comparison of expert ratings and self-assessments of Situation Awareness during a combat fatigue course. *Military Psychology, 23*(2), 125–136. <https://doi.org/10.1080/08995605.2011.550222>
- McCracken, J. H. & Aldrich, T. B. (1984). Analyses of selected LHX mission functions: Implications for operator workload and system automation goals. *Technical note ASI, 479-024*.
- McGuinness, B. (2004, June). Quantitative analysis of situational awareness (QUASA): Applying signal detection theory to true/false probes and self-ratings. In *Command and Control Research and Technology Symposium June* (S. 15-17).
- McGuinness, B. & Ebbage, L. (2002). Assessing human factors in command and control: Workload and situational awareness metrics. In *Proceedings of the 2002 Command and Control Research and Technology Symposium*. Monterey, CA, 2002.
- McGuinness, B. & Foy, L. (2000). A subjective measure of SA: the Crew Awareness Rating Scale (CARS). In *Proceedings of the first human performance, situation awareness, and automation conference* (Bd. 16, S. 286–291).
- Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. New York: Plenum Press.
- Miller, S. (2001). *Literature review: Workload measures* (Document ID: N01-006). City of Iowa. University of Iowa.
- Mitchell, D. K. (2000). *Mental workload and ARL workload modeling tools*. Army Research Lab Aberdeen Proving Ground MD.
- Mohr, G., Rigotti, T. & Müller, A. (2007). *Irritations-Skala zur Erfassung arbeitsbezogener Beanspruchungsfolgen*. Göttingen: Hogrefe.

- Muñiz, E. J., Stout, R. J., Bowers, C. A. & Salas, E. (1998). A methodology for measuring team Situational Awareness: Situational Awareness Linked Indicators Adapted to Novel Tasks (SALI-ANT). *RTO Meeting Proceedings 4*, 11. <https://apps.dtic.mil/dtic/tr/fulltext/u2/a367376.pdf#page=92>
- Neff, K. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250.
- Newman, P., Farmer, E. W. & Belyavin, A. J. (1996). Performance and Workload measurement in simulation-based training. In *AGARD conference proceedings* (Bd. 588, S. 13).
- Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors*, 33(1), 17–33.
- Oberfeld, D. (2023). Kognitive Anstrengung: kognitionspsychologische Grundlagen und Konzepte. In J. Grenzebach (Ed.) *Workshop zur Quantifizierung mentaler Beanspruchung aus Tätigkeit und Umgebungsfaktoren der Arbeit* (pp 62 – 67). https://www.baua.de/DE/Angebote/Publikationen/Fokus/Quantifizierung-mentaler-Beanspruchung.pdf?__blob=publicationFile&v=4
- O'Donnell, R. D. & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. Thomas (Hg.), *Handbook of perception and human performance: Volume II. Cognitive processes and performance* (42/1–42/49). Wiley.
- Olson, J. M. (2003). *Tactile display landing safety and precision improvements for the space shuttle* [Doctoral Dissertation]. Auburn University, Alabama, USA. <https://www.proquest.com/docview/305341427>
- Park, P. & Cha, D. W. (1998). Comparison of subjective mental workload assessment techniques for the evaluation of in-vehicle navigation systems usability. In *5th World Congress on Intelligent Transport Systems: 12-16 October 1998, Seoul, Korea: Proceedings: Toward the new horizon together for better living with ITS*. ITS America.
- Pauzié, A. (2008). A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems*, 2(4), 315. <https://doi.org/10.1049/iet-its:20080023>
- Pauzié, A. & Pachiardi, G. (1997). Subjective evaluation of the mental workload in the driving context. In T. Rothengatter & E. Carbonell Vaya (Hg.), *Traffic & transport psychology: Theory and application* (S. 172–182). Pergamon/Elsevier Science Inc.
- Pearlin, L. I. & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, 19(1), 2-21.
- Pickup, L., Wilson, J. R., Norris, B. J., Mitchell, L. & Morrisroe, G. (2005). The Integrated Work-load Scale (IWS): A new self-report tool to assess railway signaller workload. *Applied Ergonomics*, 36(6), 681–693. <https://doi.org/10.1016/j.apergo.2005.05.004>
- Poppe, P., Stiensmeier-Pelster, J. & Pelster, A. (2005). *Attributionsstilfragebogen für Erwachsene: ASF-E*. Göttingen: Hogrefe.
- Rehmann, A. J. (1995). *Handbook of human performance measures and crew requirements for flight research* (DOT/FAA/CT-TN 95/49). Crew System Ergonomics Information Analysis Center Wright-Patterson AFB OH.
- Rehmann, J. T., Stein, E. & Rosenberg, B. L. (1983). Subjective pilot Workload assessment. *Human Factors*, 25(3), 297–301. <https://doi.org/10.1177/001872088302500306>

- Reichert, M. & Perrez, M. (1993). *Fragebogen zum Umgang mit Belastungen im Verlauf (UBV). Testmappe mit Handanweisung und Testmaterial*. Bern: Huber.
- Reid, G. B. & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A scaling procedure for measuring mental workload. In M. Hancock (Hg.), *Advances in psychology: Human mental workload* (Bd. 52, S. 185–218). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)62387-0](https://doi.org/10.1016/S0166-4115(08)62387-0)
- Reid, G. B., Shingledecker, C. A., Nygren, T. E. & Eggemeier, F. T. (1981). Development of multidimensional subjective measures of workload. In *IEEE 1981 proceedings of the International Conference on Cybernetics and Society* (S. 403–406).
- Röckel, C., & Hecht, H. (2023). Regular looks out the window do not maintain situation awareness in highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 98, 368–381. doi.org/10.1016/j.trf.2023.09.015
- Rotter, J. B. (1966). Generalized Expectations for Internal versus External Control of Reinforcement. In *Psychological Monographs*, 80(1), 1-28.
- Rubio, S., Díaz, E., Martín, J. & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, 53(1), 61–86.
- Rüesch, P., Bartlomé, P. & Huber, C. (2006). *Evaluation von Messinstrumenten für den strategischen Bereich »psychische Gesundheit/Stress«*. Winterthur: Zürcher Hochschule für Angewandte Wissenschaften [ZHAW].
- Salmon, P., Stanton, N., 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>
- Schaarschmidt, U. & Fischer, A.W. (2003). *Arbeitsbezogenes Verhalten und Erlebensmuster AVEM* (2., überarbeitete Auflage). Frankfurt a. M.: Swets & Zeitlinger.
- Schaarschmidt, U., & Fischer, A. W. (2008). *Arbeitsbezogenes Verhaltens- und Erlebensmuster: AVEM (Standardform), AVEM-44 (Kurzform)*. Pearson.
- Schlotz, W., Yim, I. S., Zoccola, P. M., Jansen, L. & Schulz, P. (2011). The perceived stress reactivity scale: Measurement invariance, stability, and validity in three countries. *Psychological Assessment*, 23(1), 80–94.
- Schneider, E. E., Schönfelder, S., Domke-Wolf, M. & Wessa, M. (2020). Measuring stress in clinical and nonclinical subjects using a German adaptation of the perceived stress scale. *International Journal of Clinical and Health Psychology*, 20(2), 173– 181.
- Schulz, P., Jansen, L. J. & Schlotz, W. (2005). Stressreaktivität: Theoretisches Konzept und Messung [Stress reactivity: Theoretical concept and measurement]. *Diagnostica*, 51(3), 124-133.
- Schulz, P., Schlotz, W., Becker, P. (2004). *TICS Trierer Inventar zum chronischen Stress*. Göttingen: Hogrefe.
- Schumacher, J. & Brähler, E. (2004). *Sense of Coherence Scale -Leipziger Kurzform (SOC-L9)*. Online-Datenbank Assessment. Institut für Qualitätssicherung in Prävention und Rehabilitation an der Deutschen Sporthochschule Köln (IQPR).
- Schumacher, J., Leppert, K., Gunzelmann, H., Strauß, B. & Brähler, E. (2005). Die Resilienzskala – Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Personmerkmal. *Zeitschrift für Klinische Psychiatrie und Psychotherapie*, 53(1), 16–39

- Scielzo, S., Davis, F., Riley, J., Hyatt, J., Lampton, D. & Merlo, J. (2010). Leveraging serious games and advanced training technologies for enhanced cognitive skill development. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 54(28), 2408–2412.
<https://doi.org/10.1177/154193121005402810>
- Seipel, C. (2002). Deutsche Version der Self-Control Skala. Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). <https://doi.org/10.6102/zis137>
- Seipel, C. (2014). Deutsche Version der Self-Control Skala. In D. Danner & A. Glöckner-Rist (Hrsg.), Zusammenstellung sozialwissenschaftlicher Items und Skalen. Mannheim: GESIS.
- Selye, H. (1959). Perspectives in stress research. *Perspectives in Biology and Medicine*, 2(4), 403-416.
- Snow, M. P. & Reising, J. M. (2000). Comparison of two Situation Awareness metrics: SAGAT and SA-SWORD. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44(13), 49–52.
<https://doi.org/10.1177/154193120004401313>
- Sommer, G. & Fydrich, T. (1989). *Soziale Unterstützung: Diagnostik, Konzepte, F-SOZU*. Tübingen: Deutsche Gesellschaft für Verhaltenstherapie.
- Staller, M. S. (2011). Ich schaffe das!. Training der Kompetenzerwartung im polizeilichen Handeln. *Polizeitrainer Magazin*, 14, 25-30.
- Stanton, N., Salmon, P. & Rafferty, L. A. (2013). *Human factors methods: A practical guide for engineering and design*. Ashgate Publishing, Ltd.
- Stein, E. (1985). Air traffic controller workload: An examination of workload probe (DOT/FAA/CT-TN84/24). *Technical Note No. DOT/FAA/CT-TN84/24*. Atlantic City International Airport, NJ, Federal Aviation Administration Technical Center.
- Strater, L., Riley, J., Davis, F., Procci, K., Bowers, C. A., Beidel, D., Bunnell, B., Proaps, A., Sankaranarayanan, G., Li, B., De, S. & Cao, C. G. L. (2013). Me and my VE, part 2. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), 2127–2131. <https://doi.org/10.1177/1541931213571474>
- Tattersall, A. J. & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, 39(5), 740–748.
<https://doi.org/10.1080/00140139608964495>
- Taylor, G. S., Alicia, T. J., Turpin, T. & Surana, A. (2017). Controlling multiple unmanned aircraft from a manned helicopter: The need for advanced autonomy and refined pilot-vehicle interface. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 78–82.
<https://doi.org/10.1177/1541931213601485>
- Taylor, R. M. (1990). Development of the Situational Awareness Rating Technique (SART) as a tool for aircrew systems design. *AGARD Conference Proceedings*, 478.
- Tobaruela Arnedo, G. (2015). *A framework to assess the ability of automation to deliver capacity targets in European airspace* [Doctoral Dissertation]. Imperial College London.
- Traue, H. C., Hrabal, V. & Kosarz, P. (2000). Alltags-Belastungs-Fragebogen: Zur inneren Konsistenz, Validierung und Stressdiagnostik mit dem deutschsprachigen Daily Stress Inventory. *Verhaltenstherapie und Verhaltensmedizin*, 21(1), 15-38.
- Tsang, P. S. & Velazquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, 39(3), 358–381.
<https://doi.org/10.1080/00140139608964470>

- Udris, I. (1976). *Skalierung der "Allgemeinen Zentralen Aktiviertheit" nach Bartenwerfer. [General Consciousness Activation Rating Scale]*. <http://search.ebscohost.com/login.aspx?direct=true&db=pdx&AN=PT9001750&site=ehost-live>
- Vidulich, M. A. & Hughes, E. R. (1991). Testing a subjective metric of Situation Awareness. *Proceedings of the Human Factors Society Annual Meeting*, 35(18), 1307–1311. <https://doi.org/10.1177/154193129103501812>
- Vidulich, M. A. & Tsang, P. S. (1987). Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. *Proceedings of the Human Factors Society Annual Meeting*, 31(9), 1057–1061.
- Vidulich, M. A., Ward, G. F. & Schueren, J. (1991). Using the Subjective Workload Dominance (SWORD) technique for projective workload assessment. *Human Factors*, 33(6), 677–691. <https://doi.org/10.1177/001872089103300605>
- Waag, W. L. & Houck, M. R. (1996). Development of criterion measures of situation awareness for use in operational fighter squadrons. In Advisory group for aerospace research and development Neuilly-Sur-Seine (Hg.), *Situation awareness: Limitations and enhancement in the aviation environment* (S. 1–8).
- Wagnild, G.M. & Young, H.M. (1993). Development and psychometric validation of the Resilience Scale. *Journal of Nursing Measurement*, 1(2), 165–178.
- Warr, D., Cole, H. A. & Reid, G. B. (1986). A comparative evaluation of two subjective workload measures: The Subjective Workload Assessment Technique and the Modified Cooper Harper Scale. *Proceedings of the Department of Defence Psychology Symposium*. Colorado Springs, CO.
- Wayne, L., Waag, W. L. & Houck, M. R. (1994). Tools for assessing situational awareness in an operational fighter environment. *Aviation, Space, and Environmental Medicine*, 65(5 Suppl), A13-9.
- Wierwille, W. W. & Casali, J. G. (1983). A validated rating scale for global mental workload measurement applications. *Proceedings of the Human Factors Society Annual Meeting*, 2(27), 129–133.
- Wierwille, W. W. & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(2), 263–281.
- Wierwille, W. W., Rahimi, M. & Casali, J. G. (1985). Evaluation of 16 measures of mental work-load using a simulated flight task emphasizing mediational activity. *Human Factors*, 27(5), 489–502.
- Willems, B. (2000). *Development of the Situation Awareness Verification and Analysis Tool (SA-VANT)*. unpublished. FAA William J. Hughes Technical Center.
- Willems, B. & Heiney, M. (2002). *Decision support automation research in the en route air traffic control environment*. DOT/FAA/CT-TN02/10. Federal Aviation Administration Technical Center Atlantic City, NJ.
- Williams, C. & Crawford, G. (1980). *Analysis of subjective judgment matrices (Report-2572-AF)*. Rand Corporation.
- Wolf, A. S. (2014). Einfluss von Stress auf die Kognition. *Zeitschrift für Orthomolekulare Medizin*, 12(04), 20-23.
- Xiao, Y. M., Wang, Z. M. & Lan, Y. J. (2005). The appraisal of reliability and validity of subjective workload assessment technique and NASA-task load index. *Chinese Journal of Industrial Hygiene and Occupational Diseases*, 23(3), 178–181.
- Yee, S., Nguyen, L., Green, P., Oberholtzer, J. & Miller, B. (2007). *Visual, auditory, cognitive, and psychomotor demands of real in-vehicle tasks*. University of

- Michigan, Ann Arbor, Transportation Research Institute. <https://deep-blue.lib.umich.edu/handle/2027.42/64456>
- Zakerian, S. A., Zia, G., Nasl Seraji, G., Azam, K. & Morteza pour, A. (2018). Reliability and validity of the Driver Activity Load Index for assessing mental workload among drivers in production companies. *Journal of Occupational Hygiene Engineering*, 5(2), 65–71. <https://doi.org/10.21859/johe.5.2.65>
- Zhang, T., Yang, J., Liang, N., Pitts, B. J., Prakah-Asante, K., Curry, R., Duerstock, B., Wachs, J. P., & Yu, D. (2023). Physiological Measurements of Situation Awareness: A Systematic Review. *Human Factors*, 65(5), 737–758. <https://doi.org/10.1177/0018720820969071>
- Zijlstra, F. R. H. (1993). *Efficiency in work behaviour: A design approach for modern tools* [Doctoral Dissertation]. Delft University of Technology, Soesterberg.
- Zijlstra, F. R. H. & Meijman, T. (1989). Het meten van mentale inspanning met behulp van een subjectieve methode (measurement of mental effort with a subjective method). In *Mentale Belasting en Werkstress. Een Arbeidspsychologische Benadering* (S. 42–61). Assen/Maastricht: Van.
- Zijlstra, F. R. H. & van Doorn, L. (1985). *The construction of a scale to measure perceived effort* [Technical Report]. Delft University of Technology, Netherlands.

6 Abbildungsverzeichnis

Abbildung 1	Übersicht der Faktoren des DALI und deren Beschreibung	17
Abbildung 2	Skala und Beschreibung der ISA.....	19
Abbildung 3	Farblich kodierte Skala der IWS	21
Abbildung 4	Rating-Skala der MACE.....	23
Abbildung 5	Die MCH	25
Abbildung 6	Definition der Subskalen des NASA-TLX.....	30
Abbildung 7	Skala der RSME.....	33
Abbildung 8	Dreidimensionales Workload Konstrukt der SWAT	36
Abbildung 9	Antwortbogen des SWORD	38
Abbildung 10	Darstellung der SWORD Beurteilungsmatrix und der vier Dimensionen.....	38
Abbildung 11	Verbale Anker der vier VACP Dimensionen	41
Abbildung 12	Elf-stufige Skala der VOSO	43
Abbildung 13	Multidimensionale Matrix der WCI/TE Methode	45
Abbildung 14	Ratingbogen des WP	47
Abbildung 15	Verfahren zur WL-Bestimmung laut Casner & Gore	49
Abbildung 16	Verfahren zur SL Bestimmung laut Rehmann	49
Abbildung 17	Definition der CARS Skalen	51
Abbildung 18	Definition der fünf Ratingstufen von CLSA	53
Abbildung 19	Items zur Erhebung von SA durch die MARS	56
Abbildung 20	Übersicht von Validierungsstudien zu SAGAT.....	61
Abbildung 21	Dimensionen und Items zur Erfassung der Team-SA mittels SALIANT	64
Abbildung 22	Items und Kategorien der SARS	68
Abbildung 23	Übersicht der zehn Items und Skala von SART	70
Abbildung 24	Übersicht aller 31 Items der SASRF.....	72
Abbildung 25	Beschreibung der drei Dimensionen von VESARS.....	81
Abbildung 26	Vergleich von SAGAT und SPAM	83
Abbildung 27	Items des ABF	86
Abbildung 28	Items des IR.....	91
Abbildung 29	Items des PSS	93
Abbildung 30	Items des Mental Health Fragebogens	95
Abbildung 31	Reaktionskomponenten und Variablen UBV	100
Abbildung 32	Items SCS-D.....	103
Abbildung 33	Items PSRS	109
Abbildung 34	Items GSE-3.....	113
Abbildung 35	Items SSA	115
Abbildung 36	Items der Pearlin Mastery Scale	116
Abbildung 37	Items der IE-4 Skala.....	118
Abbildung 38	Items des SCS	121
Abbildung 39	Items der H-Skala	123
Abbildung 40	Items der RS-11.....	125
Abbildung 41	Messmethoden auf Basis des transaktionalen Stressmodells	127

7 Abkürzungsverzeichnis

ATC	Air Traffic Control
ATCO	Air Traffic Control Operator
HF	Human Factors
MWL	Mental Workload
RT	Reaction Time
SA	Situational Awareness
SME	Subject matter expert
VL	Versuchsleitung
WL	Workload