

Psychophysikalische Untersuchungen des wahrgenommenen Komforts in Flugzeugkabinen

Von der Fakultät für Mathematik und Naturwissenschaften der Carl von Ossietzky Universität Oldenburg zur Erlangung des Grades und Titels eines Doktors der Naturwissenschaften (Dr. rer. nat.) angenommene Dissertation

von

Ingo Baumann geboren am 23. April 1974 in Oldenburg

Erstgutachter: Prof. Dr. Volker Mellert Zweitgutachter: Prof. Dr. Jesko L. Verhey Tag der Disputation: 04.12.2009

Inhaltsverzeichnis

Al	bbildu	ingsverzeichnis	v	
Та	Tabellenverzeichnis viii			
Al	bkürz	ungsverzeichnis	×	
Ζι	usamr	nenfassung	xiii	
1	Einle	eitung	1	
2	ICE	I – Ideal Cabin Environment	5	
	2.1	Überblick	5	
	2.2	Messkampagne	6	
	2.3	Verwendete Simulationsumgebungen	7	
	2.4	Ausgewählte Umgebungsparameter und Variationsbereich	12	
	2.5	Nicht systematisch variierte Parameter	13	
	2.6	Messwertüberwachung und -speicherung	13	
	2.7	Versuchsdesign	14	
	2.8	Teilnehmer	17	
	2.9	Fragebögen	18	
	2.10	Messprogramm	20	
	2.11	Tagesablauf	20	
	2.12	Datenerhebung und -speicherung	24	
3	ICE	II – Auswertung	25	
	3.1	Vorbemerkung	25	
	3.2	Psychologische Analyse mittels Fragebögen	26	
	3.3	Ratingskalen	27	
	3.4	Normierung	28	
	3.5	Rücktransformation	30	
		3.5.1 Skalenproblematik (Rück-)Transformation	31	
	3.6	Zuordnungsvorschrift ϕ	33	

		3.6.1	Quellskala	33
		3.6.2	Zielskala	33
		3.6.3	Grenzen	33
		3.6.4	Zuordnung	34
	3.7	Beispie	eldarstellung	36
	3.8	Zusam	menfassung und weiteres Vorgehen	37
4	ICE	III – Di	istribution of Subjective Assessments in a Controlled Aircraft	
	Env	ironme	nt	38
	4.1	Abstra	act	38
	4.2	Introd	uction	39
	4.3	Data o	collection	40
	4.4	Data a	analysis	41
		4.4.1	The Poisson distribution	43
		4.4.2	Required raw data for further analysis	43
		4.4.3	Fitting the observed distribution with a Poisson distribution .	44
		4.4.4	Validation of approximation by applying a goodness-of-fit test	45
		4.4.5	Alternative method for approximation: minimizing d_{max}	46
	4.5	Result	s of Kolmogorov-Smirnov goodness-of-fit test (KS-test)	46
		4.5.1	Example	46
		4.5.2	Analysis of sections	47
		4.5.3	Variance-to-mean ratio (VMR)	49
		4.5.4	Results per section	51
		4.5.5	Summary	54
	4.6	Applic	eation: Predicted Percentage of Dissatisfied - PPD	56
		4.6.1	Definition	56
		4.6.2	Quality of PPD prediction	56
	4.7	Conclu	sion and outlook	57
	4.8	Acknor	wledgement	58
	4.9	Appen	dix	60
5	ICE	IV – Ir	nvestigation of temporary threshold shifts in aircraft cabins	61
	5.1	Introd	uction and theoretical background	61
	5.2	Develo	opment and testing of a fast audiometric tool for threshold dif-	
		ference	e measurements	67
		5.2.1	Implementation of the audiometric tool	68
		5.2.2	Stimuli and presentation	68
		5.2.3	$Measurement \ protocols \ . \ . \ . \ . \ . \ . \ . \ . \ . \ $	70

	5.3	Heari	ng threshold fluctuations in a non-TTS causing environment \therefore	. 72
		5.3.1	Question	. 72
		5.3.2	Exposure	. 73
		5.3.3	Participants	. 73
		5.3.4	Procedure	. 73
		5.3.5	Results	. 73
		5.3.6	Summary of hearing threshold variation without TTS	. 86
	5.4	Heari	ng threshold differences after 7 hours at 80 $\mathrm{dB}(\mathrm{A})$ exposure	. 87
		5.4.1	Question	. 87
		5.4.2	Exposure	. 87
		5.4.3	Participants	. 87
		5.4.4	Procedure	. 88
		5.4.5	Results	. 89
		5.4.6	Summary of hearing threshold variation in TTS-causing envi-	
			ronment	. 93
	5.5	Thres	hold measurements before and after simulated flights \ldots .	. 93
		5.5.1	Question	. 93
		5.5.2	Exposure	. 93
		5.5.3	Participants	. 94
		5.5.4	Procedure	. 95
		5.5.5	Results	. 96
		5.5.6	Summary of hearing threshold variations before and after si-	
			mulated flights	. 99
	5.6	Summ	hary of results	. 100
	5.7	Concl	usion \ldots	. 101
	5.8	Ackno	owledgement	. 103
	5.9	Appe	ndix	. 104
6	ЦΕ		Impact of sound and vibration on health travel comfort an	J
0		ACE -	impact of sound and vibration on health, travel connort and	u 100
	6 1	Abstr	ace of crew	109
	0.1 6 9	Intro	act	1109
	0.2 6.3	Tost	lucuon	110
	0.0	Lest (Promiso	110
		629	Questionnaire design	111
		0.J.Z	Flight tost	112
		0.0.0	Simulator tests	110 110
		0.0.4	Dimutatol legis	. 110

Inhalts verzeichnis

	6.4	Data analysis	. 119
		6.4.1 Data from flight tests	. 121
		6.4.2 Data from simulator tests	. 126
	6.5	Conclusion	. 127
7	Zus	ammenfassung und Diskussion	129
Literaturverzeichnis 14			147
Stichwortverzeichnis 1			154
Publikationsliste 156			156
Le	Lebenslauf 1		

Abbildungsverzeichnis

1.1	Sich verändernde Umweltparameter während eines Kurzstreckenfluges	2
2.1	ACE - Facility, Funktionsskizze	9
2.2	Aufbau des FTF	10
2.3	Innenansichten des Flight Test Facility (FTF) - Simulators	11
2.4	Bestuhlung des FTF-Simulators	12
2.5	Messfühler - Fotos	14
2.6	Voll faktorielles Versuchsdesign (3 Parameter)	15
2.7	Voll faktorielles Versuchsdesign (4 Parameter)	15
2.8	Finales Versuchsdesign	17
2.9	Altersverteilung der Probanden im ICE-Versuch	18
2.10	Ablaufplan der vorgestellten Messkampagne	20
3.1	Beispiel 1 für ein mögliches Antwortverhalten	26
3.2	Beispiel 2 für ein mögliches Antwortverhalten	27
3.3	Beispiel für ein mögliches Antwortverhalten – 4er und 7er Skala \ldots .	29
3.4	Beispiel für ein mögliches Antwortverhalten – umskaliert	29
3.5	Beispiel einer Häufigkeitsverteilung mit 40 Teilnehmern, normiert und	
	umskaliert auf den Antwortbereich '0' bis '1'	30
3.6	Ablaufskizze zur Verarbeitung der Fragebogendaten	32
3.7	Antwort x_k mit den Grenzen $x_k \stackrel{[\leftarrow}{} und x_k \stackrel{\rightarrow]}{} \dots \dots \dots \dots \dots$	34
3.8	Antwort y_l mit den Grenzen $y_l \stackrel{[\leftarrow}{=} \text{ und } y_l \stackrel{\rightarrow}{\rightarrow}$	34
3.9	Antwort x_k zugeordnet \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots	35
3.10	Umsetzung der Zuordnungstabelle	36
3.11	Ergebnis der Umskalierung	37
4.1	Example of an observed answer distribution for Mean-sym-pain	42
4.2	Example of an observed answer distribution for C-Cond	47
4.3	Frequency in % of probabilities p_{KS} with $p_{KS} \ge 95\%$	48
4.4	Histogram of calculated variance-to-mean ratios (VMR)	49
4.5	Dependency between calculated VMR of observed data and frequency	50
4.6	Result of section "behavior" $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	51

4.7	Results of section "environmental perception"	52
4.8	Results of "comfort" section $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	53
4.9	Results of section "personal mental status" $\ldots \ldots \ldots \ldots \ldots$	53
4.10	Results of section "perceived symptoms"	54
4.11	Frequency in % of probabilities p_{KS} with $p_{KS} \ge 95\%$ for all investi-	
	gated questionnaire sections	55
4.12	Predicted and observed PPD as scatter-plot	57
5.1	TTS_4 -values as a function of exposure duration	63
5.2	Hypothetical growth as a prediction of a TTS_2	64
5.3	Different regression functions as forecast for recovery	65
5.4	Schematically cut of modified Békésy-procedure	67
5.5	Set-up of audiometric tool	68
5.6	Scheme of three pulse levels of pulse train	69
5.7	Different measurement protocols HTM1 and HTM2 \ldots	71
5.8	Schematic diagram of used frequency blocks FB1 to FB3 \ldots .	71
5.9	Schematic plot of 50 dB(A) test $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	74
5.10	Individual hearing thresholds L_T measured six times on three days	75
5.11	Hearing threshold differences $\Delta L_T(f,p)$	77
5.12	Hearing threshold differences	80
5.13	Standard deviations $\sigma_{\langle d,P \rangle}$ and $\sigma_{\langle tod,P \rangle}$	82
5.14	Standard deviations σ for all participants	83
5.15	Standard deviations $\sigma_{\langle f, P \rangle}$ for all frequencies	84
5.16	Averaged time to perform a hearing threshold measurement	86
5.17	Third-octave octave spectra of A320 sound file	88
5.18	Schematic plot of timeline during the 7 hour / 80 dB(A) exposure $\ .$.	89
5.19	$\Delta L_{T,2min}, \Delta L_{T,15min}$ and $\Delta L_{T,30min}$ for different participants	90
5.20	$\Delta L_{T,2min}, \Delta L_{T,15min}$ and $\Delta L_{T,30min}$ for different frequencies	91
5.21	Standard deviations for all frequencies	92
5.22	Third-octave spectra (74 dB(A), 6^{th} row) in A310 simulator \ldots	94
5.23	Schematic plot of timeline during the tests within the ICE-project	95
5.24	Differences in hearing thresholds $(\Delta L_{T,30min})$ for 88 passengers	97
5.25	Standard deviations for the participants in one day's flight	99
5.26	Instruction to participant	104
5.27	Averaged level differences ΔL_{rev} between reversal points	106
5.28	Averaged level differences ΔL_{rev} between the reversal points	107
5.29	Age distribution	108

6.1	Distribution of questionnaires to "test passengers" in the ACE simulator 113
6.2	Distribution of A- vs. B-weighted sound level in cockpit and cabin 115
6.3	dB(A)-levels in cockpit and cabin along aircrafts A330 and A340 116
6.4	Increase of dB(B) level over dB(A) level $\ldots \ldots \ldots \ldots \ldots \ldots \ldots 116$
6.5	Weighted vibration levels in dB
6.6	"Wired" pilot during pre-test in the cockpit
6.7	Plane of [Factor 1 – Factor 3] of an 11-dimensional space of perception. 122
6.8	Plane of [Factor 2 – Factor 4] of an 11-dimensional space of perception. 123
6.9	ANOVA of development of subjective air quality
6.10	ANOVA of symptom "swollen feet" $\ldots \ldots \ldots$
6.11	Scheme to illustrate the trade-off $\ldots \ldots $
7.1	Häufigkeiten von statistisch signifikant angepassten Verteilungen 131
7.2	Gemessene und vorhergesagte Antwortverteilung für 'E-passenger' 132
7.3	Beobachtete und approximierte Antwortverteilung für 'Mean-sym
	treeze'
7.4	Antwortverhalten von 10 Probanden bei unterschiedlichen Tempera-
	turen
7.5	$\lambda\text{-}Werte$ sowie PPD-Werte für verschiedene Temperaturbedingungen . 136
7.6	$\lambda\text{-}Werte$ sowie PPD für verschiedene Schallbedingungen
7.7	${\rm Logarithmic \ fit \ of \ restituion \ curves} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 140$
7.8	Darstellung verschiedener Parameterkombinationen a und b
7.9	Ergebnis aller TTS_{30} Untersuchungen
7.10	Darstellung der Ergebnisse verschiedener Hörschwellenverschiebungs-
	experimente

Tabellenverzeichnis

2.1	Messplan für die FTF-Tests
4.1	Overview of the 5 analyzed sections
4.2	Table of Confusion 45
4.3	Overview of items and indices
5.1	Summary of used measurement protocol HTM1
5.2	Summary of the observed (sample) standard deviations
5.3	Averaged observed threshold differences $\Delta L_{T,30min}$ for simulated flights 100
5.4	Observed threshold differences for 2, 15 and 30 minutes after exposure 101
6.1	Test design in ACE simulator

Abkürzungsverzeichnis

ACE	Aircraft Cabin Environment
ANN	Artificial neural network
AVI	Avitronics Research, Griechenland
BRE	Building Research Establishment Ltd., England
CAA	Civil Aviation Authority, England
DLR	Deutsche Forschungsanstalt für Luft- und Raumfahrt e.V.
DVT	deep vein thrombosis
FACE	Friendly Aircraft Cabin
FTF	Flight Test Facility
HEACE	Health Effects in Aircraft Cabin Environment
IBP	Fraunhofer-Institut für Bauphysik
ICE	Ideal Cabin Environment
MUV	Medizinische Universität Wien, Österreich
NLR	Stichting National Lucht en Ruimtevaartlaboratorium, Niederlande
PCA	Principle Component Analysis
PDA	Personal Digital Assistant
PDM	Predictive Design Model
SVOC	semi-volatile organic components - mittel-flüchtige organische
	Komponenten
UCL	University College London, England

 ${\sf VOC}$ \quad volatile organic components - flüchtige organische Komponenten

Zusammenfassung

Hauptziel des Forschungsprojekts ICE¹, dokumentiert in Kapitel 2, war die Erstellung eines neuartigen Vorhersagemodells, welches die Wahrnehmung und die Zufriedenheit der Passagiere mit den systematisch variierten Umweltparametern Temperatur, Luftfeuchtigkeit, Druck und Akustik vorhersagt und Empfehlungen für zukünftige Kabinengestaltungen in Form eines neuen europäischen Standards gibt. Dazu wurden Interaktionen der genannten Umgebungsparameter auf Gesundheit und Komfort bei ca. 1.500 Passagieren untersucht, die an 7-stündigen simulierten Flügen teilgenommen haben.

Nach einer notwendigen Anpassung der aus der Vorverarbeitung gelieferten Fragebogendaten (Transformation teils kontinuierlicher Skalen in diskrete Intervallskalen, Kapitel 3) wurde ein neues Verfahren zur Parametrisierung beobachteter Antwortverteilungen entwickelt und angewendet (Kapitel 4). Die Anpassung der Verteilungen erfolgt mittels der (diskreten) Poisson-Verteilungsfunktion, die durch den Parameter λ vollständig beschrieben ist. Für den analysierten Datensatz (Antworten auf unipolaren Skalen bei einem Varianz-zu-Mittelwert Verhältnis (VMR) von $R \geq 0,72$) wurde für 95% der beobachteten Antwortverteilungen eine statistisch signifikant gute Anpassung erreicht. Damit wird eine erhebliche Datenreduzierung unter Beibehaltung der Detailinformationen zum Antwortverhalten befragter Passagiere ermöglicht. Es erlaubt weiter eine Interpolation für Wertebereiche, die nicht direkt gemessen wurden.

Im Anschluss wird die Frage untersucht, ob durch anhaltende Exposition durch Kabineninnengeräusche eine zeitweilige Schädigung des Gehörs (Temporary Threshold Shift - TTS) auftreten kann (Kapitel 5). Dazu wurde zunächst mit Hilfe eines neu entwickelten Messsystems die Hörschwellenfluktuation innerhalb eines Tages durch Bestimmung der individuellen Ruhehörschwelle jeweils morgens und abends bei 8 Probanden an drei verschiedenen Tagen bestimmt. Bei einem mittleren Expositionspegel von 50 dB(A) war von keiner lärminduzierten Hörschwellenverschiebung auszugehen, so dass anhand dieser Messungen die Möglichkeit bestand, Tagesschwankungen der Hörschwelle von denen einer möglichen lärmbedingten Verschiebung zu unterscheiden. Es ergab sich eine (nicht signifikante) Verschiebung der

¹Ideal Cabin Environment, Laufzeit 2005 - 2009

Hörschwelle um 0, 30 dB \pm 0, 19 dB. In einem zweiten Versuch wurde bei 6 Probanden die Hörschwelle vor und zu drei verschiedenen Zeiten nach einer 7-stündigen Exposition mit Kabineninnengeräuschen bei einem Pegel von 80 dB(A) gemessen. Es ergaben sich sowohl 2 min nach Ende der Exposition eine signifikante Verschiebung von 3, 90 dB \pm 0, 51 dB zur Hörschwellenmessung am Morgen als auch nach 15 min mit 1, 66 dB \pm 0, 50 dB. Für 30 min mit 0, 29 dB \pm 0, 48 dB ist die Verschiebung nicht signifikant unterschiedlich zur Messung am Morgen. Der beobachtete Erholungsverlauf entsprach dem in der Literatur erwähnten logarithmischen Restitutionsverlauf. Die Untersuchung der Hörschwelle vor und 30 min nach einer 7-stündigen Exposition mit Kabineninnengeräuschen (74 dB(A)) im Rahmen des ICE-Simulatorversuchs ergab eine signifikante Verschiebung von 0, 72 dB \pm 0, 12 dB und damit eine (leichte) temporäre Hörschwellenverschiebung (TTS).

Abschließend werden Auswirkungen von Schall und Vibrationen auf Gesundheit, Reisekomfort und Arbeitsleistung der Flugbegleiter und Piloten während simulierter als auch während realer Langstreckenflüge als ein Ergebnis des HEACE-Projekts² dargestellt (Kapitel 6). Messungen des Schallfeldes an Bord von Airbus A330 und A340 Flugzeugen sowie Ergebnisse der Fragebogendatenanalyse werden gezeigt. Dabei ist der Lautstärkepegel neben der Luftqualität ein entscheidender Parameter, der sich negativ auf die subjektive Befindlichkeit auswirkt und einen signifikanten Einfluss auf die Arbeitsleistung der Crew ausübt, speziell bei langer Einwirkdauer.

Summary

Within the research project ICE^3 , a new predictive design model was developed to identify the impact of environmental conditions (like temperature, humdity, pressure, and noise) on passenger comfort and well-being and to give recommendations for future cabin design in an European standard. Therefore, 1500 passengers took part in several simulated flights of seven hours duration, in which their reaction was monitored.

After a necessary adjustment of preprocessed questionnaire data (transformation of (continous) scales into discrete intervall scales, chapter 3), a new method to parameterize observed answer distribution was developed and used (chapter 4). It is done by an approximation with (discrete) Poisson-distribution, characterized completely by a single parameter λ . For subjective ratings on a unipolar scale and a calculated

 $^{^2}$ Health Effects in Aircraft Cabin Environment, europäisches Forschungsprojekt, 2001 – 2005 3 Ideal Cabin Environment, running from 2005 to 2009

variance-to-mean ratio of $R \ge 0.72$ in raw distribution data, a success rate of 95% in representing the observed data distribution was identified. This approach allows for immense data reduction and a detailed representation of an average participant's impression (answer distributions of questionnaire data) of a defined environmental condition at the same time and enables for a prediction of not measured, interim situations.

The aim of chapter 5 is the investigation of potential temporary threshold differences during long-haul flights due to noise load in aircraft cabin. Therefore, a precise, effective and portable measurement system for hearing threshold measurement in field has been developed, extensively tested and validated. In order to be able to identify possible temporary hearing threshold shifts (TTS) in an aircraft cabin, the statistics of normal hearing threshold variation within a day have been tested in a non-TTS causing environment (50 dB(A)) in order to be able to discriminate statistical fluctuations from bias that may be identified as TTS. A threshold shift of 0.30 dB \pm 0.19 dB (statistically not significant, t test, p < 0.05) was measured. A second acoustic environment consisted of noise, spectrally shaped as aircraft cabin interior noise, at a level of 80 dB(A) was presented to 6 participants during 7 hours. Hearing thresholds were measured before and three times after exposure. A statistically significant threshold difference 2 min and 15 min after exposure of $3.90 \text{ dB} \pm 0.51 \text{ dB}$ and $1.66 \text{ dB} \pm 0.50 \text{ dB}$ were observed. Measurements 30 min after exposure $(0.29 \text{ dB} \pm 0.48 \text{ dB})$ showed no significant difference anymore to the measurement before. The logarithmic characteristics of observed restitution was in accordance with data from literature. After 7 hours simulated flights at an aircraft cabin noise exposure level of 74 dB(A), a statistically significant threshold shift of $0.72 \text{ dB} \pm 0.12 \text{ dB}$ and therefore a (slight) TTS was observed.

Finally, the impact of sound and vibration on health, travel comfort and performance of flight attendants and pilots during simulated as well as real long-haul-flights are presented as one result of the HEACE-project⁴ (chapter 6). Measurements of sound field distribution of Airbus A330 and A340 aircrafts and results of questionnaire analyses are shown. As a result, sound level and air quality are important parameters with negative impact on subjective feeling and significant bias on the crew's performance, in particular when effective during long periods.

 $^{^4\}mathrm{Health}$ Effects in Aircraft Cabin Environment, European research projekt, 2001 – 2005

1 Einleitung

Das (Fern-) Reiseaufkommen ist in den vergangenen Jahrzehnten stetig gestiegen. Im Jahr 2007 wählten bei Freizeitreisen ins Ausland 40,8% das Flugzeug (PKW: 44,6%, Bus: 9,2%, Bahn: 3,9%). Bei Geschäftsreisen lag der Anteil der Flugreisen bei 62,7% und damit deutlich vor PKW (29,6%), Bahn (3,8%) oder Bus $(3,1\%)^1$. Zudem ist u.a. aufgrund der fortschreitenden Globalisierung zu erwarten, dass eine zunehmende Anzahl von Passagieren immer mehr Zeit in Flugzeugen verbringen wird [1]. So stieg beispielsweise die Zahl der abgefertigten Passagiere auf den Flughäfen weltweit von 2 Mrd. (1991) auf 3.5 Mrd im Jahr 2003. Dies entspricht einem jährlichen Wachstum von durchschnittlich 4,8%. In Europa wuchs das Passagieraufkommen im gleichen Zeitraum um durchschnittlich 5,8% [2]. Prognosen für das Jahr 2020 rechnen mit vergleichbaren Zuwachsraten [1]. Dabei gilt sowohl für Urlaubs- als auch für Dienstreisen das gleiche Ziel: Ein entspanntes Reisen, frei von möglichen Gesundheitsbeeinträchtigungen.

Fluggesellschaften erlauben durch den Einsatz von Großraumflugzeugen (sog. *widebody aircraft*) mit bis zu 10 Sitzen in einer Reihe (3-4-3 Bestuhlung) im Rahmen von Lang- und Ultralangstreckenflügen auch die Überwindung großer und größter Entfernungen. Während des Aufenthalts an Bord befindet sich der Passagier in einer komplexen Umgebung mit multimodaler Wahrnehmung. Er sieht sich dabei Umgebungsbedingungen ausgesetzt, die an Bord des Flugzeuges von verschiedenen Systemen künstlich erzeugt werden. Eine Auswahl sich während eines Fluges verändernder Umweltparameter ist in Abbildung 1.1 dargestellt [3].

In ihrer Summe unterliegt der menschliche Körper, dessen Physiologie auf die Existenz auf Meereshöhe optimiert ist [4], dabei zahlreichen externen Einflussfaktoren. Von Interesse sind neben den direkten und unmittelbaren Auswirkungen dieser Faktoren auf die Komfortwahrnehmung auch die gesundheitlichen Beeinträchtigungen durch potentielle Langzeitschäden, z.B. durch die Entstehung von Thrombosen (*deep vein thrombosis (DVT*)) durch (zu) langes Sitzen [5]. Ebenso stellt sich die

¹Pressemitteilung Nr. 364 des Statistischen Bundesamt Deutschland (24.09.2008)



Abbildung 1.1: Übersicht über sich verändernde Umweltparameter während eines realen Fluges (Flugdauer ca. 90 Minuten): Druck, Luftfeuchtigkeit, Temperatur, CO₂- und CO-Gehalt und Schall (L_{A,eq}) [3].

Frage, ob und inwieweit einzelne Parameter untereinander interagieren und ob es synergetische Effekte gibt.

Sowohl Fluggesellschaften als auch die Flugzeughersteller selbst legen zunehmend Wert darauf, Kabineninnenausstattungen (das *Lining*) zu entwickeln, die größtmöglichen Komfort anbieten unter gleichzeitiger Berücksichtigung der Herstellungskosten und der laufenden Betriebskosten. Detaillierte Kenntnisse über die Wahrnehmung, die Beurteilung und somit der Akzeptanz der Passagiere bzgl. der herrschenden Umweltbedingungen sind dabei als entscheidender Wettbewerbsvorteil anzusehen.

Um diese Abhängigkeiten zu identifizieren, sind reproduzierbare Untersuchungen mit systematisch veränderbaren Umweltbedingungen an Bord von Flugzeugkabinen notwendig. Und gerade hier liegt die Schwierigkeit, weil dieses während realer Flüge nahezu unmöglich ist. Ebenso haben bisherige Untersuchungen immer nur einzelne Teilaspekte beleuchtet, wie z.B. die (kombinierte) Wirkung zwischen Akustik und Vibrationen oder zwischen Temperatur und anderen Klimaparametern.

Mit den europäischen Forschungsprojekten CabinAir [6] und FACE [7] sowie HEACE [8] wurden in der Vergangenheit erstmalig relevante Parameter für den Passagierkomfort (*ride-comfort*) sowie für die Arbeitsbelastungen (HEACE) an Bord identifiziert und detailliert kombiniert untersucht. Dabei stand speziell das Zusammenwirken von Klima (Temperatur), Schall und Vibrationen im Vordergrund.

Wie aus Abbildung 1.1 ersichtlich wird, wirken jedoch noch weitere Umweltparameter auf die Flugpassagiere und die Crew ein. Daher wurden im Rahmen des $3\frac{1}{2}$ Jahre andauernden Forschungsprojekts Ideal Cabin Environment (ICE) die Interaktion weiterer ausgewählter Umgebungsparameter (wie z.B. der Luftdruck) auf die Gesundheit und den Komfort von ca. 1.500 Passagieren untersucht. Dabei wurden detailliert gestaltete Simulationsumgebungen verwendet, die neben gesundheitlichen Aspekten auch Komfortuntersuchungen ermöglichen. Hauptziel des Projekts war die Erstellung eines neuartigen Vorhersagemodells, das *Predictive Design Model* (PDM), welches es ermöglicht, die Wahrnehmung und vor allem die Zufriedenheit der Passagiere mit den getesteten Umweltbedingungen vorherzusagen und damit Empfehlungen für die zukünftige Kabinengestaltung geben zu können.

Die Arbeitsgruppe Akustik der Universität Oldenburg war in diesem Projekt zum einen verantwortlich für die Datenanalyse und Aufbereitung der gesammelten und bereits vorverarbeiteten Fragebogendaten hinsichtlich der Modellierung. Zum anderen wurden gesundheitliche Auswirkungen auf das Gehör (mögliche Verschiebung der Ruhehörschwelle nach einem mehrstündigen Flug) untersucht. Die vorliegende Arbeit stellt im Folgenden einige dieser Forschungstätigkeiten und -ergebnisse umfassend vor.

Beginnend mit **Kapitel 2** wird zunächst ausführlich das europäische Forschungsprojekt ICE dokumentiert. Die im Rahmen dieses Projekts durchgeführte Untersuchung lieferte die Daten (und Rahmenbedingungen) für die weiteren beschriebenen Auswertungsschritte.

Im Bereich der Fragebogenanalyse wurde im Laufe der Auswertung ein neuartiges Verfahren zur Parametrisierung von beobachteten Antwortverteilungen entwickelt, um später die in diesem Projekt erhobenen Fragebogendaten komprimiert auf einfache Art und Weise in das Vorhersagemodell überführen zu können.

Dazu gehört zunächst die Anpassung der aus der Vorverarbeitung gelieferten Fragebogendaten (**Kapitel 3**) sowie die anschließende Parametrisierung der Antwortverteilungen mittels Approximation mit einer diskreten Verteilungsfunktion (Poisson-Verteilung) in **Kapitel 4**.

Das sich anschließende **Kapitel 5** befasst sich mit einem gesundheitlichen Aspekt des Gehörs und untersucht, ob sich während einer 7-stündigen Exposition mit typischen Umgebungsparametern, wie man sie in der Flugzeugkabine vorfindet, eine gesundheitliche Beeinträchtigung des Hörvermögens in Form einer kurzzeitigen Hörschwellenverschiebung einstellen kann.

Neben den Untersuchungen des ICE-Projekts, bei denen die Wirkung von Umgebungsparametern auf den Komfort, das Wohlbefinden und den allgemeinen Gesundheitszustand von Passagieren im Vordergrund stand, wurden im Rahmen des HEACE-Projekts die Effekte physikalischer Umgebungsparameter, insbesondere Schall und Vibrationen, auf Gesundheit, Komfort und Arbeitsleistung der Crew am Arbeitsplatz 'Kabine' untersucht. Ergebnisse, die im Journal "Aerospace Science and Technology" veröffentlicht² wurden, sind in **Kapitel 6** dargestellt.

In **Kapitel 7** folgt abschließend die Zusammenfassung und eine ausführliche Diskussion aller präsentierten Ergebnisse sowie ein Fazit auf Seite 146.

²Aerospace Science and Technology, Band 12, Ausgabe 1, Januar 2008, Seiten 18-25

2 ICE I – Ideal Cabin Environment

Das folgende Kapitel befasst sich mit dem europäischen Forschungsprojekt ICE. Nach einer allgemeinen Projektvorstellung werden die im Projekt verwendeten Simulationssysteme näher vorgestellt. Neben Messdesign und Versuchsablauf werden die weiteren Auswertungsschritte hin zum Projektziel von ICE – der Erstellung eines Vorhersagemodells für den Passagier-Komfort in der Flugzeugkabine – beschrieben. Daraus leiten sich grundlegend die Rahmenbedingungen für die weiteren Auswertungsschritte ab, die sich in den Folgekapiteln anschließen.

2.1 Überblick

ICE ist ein multinationales europäisches Forschungsprojekt, welches im 6. Rahmenprogramm von der EU gefördert¹ wurde. Zwischen Oktober 2005 und März 2009 wurden die Auswirkungen der in einer Flugzeugkabine herrschenden Umgebungsbedingungen auf die Gesundheit und das Wohlbefinden der Passagiere untersucht. Das ICE-Konsortium, bestehend aus 15 Partnern aus acht europäischen Ländern, befasste sich dabei mit der Untersuchung von komfortrelevanten Parameter und ihrer Kombination zum Zwecke der Erstellung von komfortoptimierten Empfehlungen in Form eines Vorhersagemodells [5]. Gesundheitsrelevante Aspekte wie die Auswirkungen auf das Herz-Kreislauf-System sowie die Lungenfunktion wurden dabei ebenfalls mit untersucht. Eine Liste der teilnehmenden Partner im ICE-Projekt befindet sich in Kapitel 4.8, Seite 58.

Teile dieser Dissertation sind im Rahmen des ICE-Projekts erarbeitet worden. Die Arbeitsgruppe Akustik der Universität Oldenburg war für die detallierte Analyse der gelieferten vorverarbeiteten Fragebogendaten und die Ableitung von Vorhersagemodellen zur Implementierung in das Gesamtmodell verantwortlich.

¹Vertragsnummer: AST4-CT-2005-516131

2.2 Messkampagne

Das im Rahmen des Projekts entwickelte Vorhersagemodell (Predictive Design Model (PDM)) beruht auf den Ergebnissen von simulierten Flügen in bodengebundenen Kabinensimulatoren. An insgesamt 35 Flügen, in denen systematisch eine Auswahl von Umgebungsbedingungen tageweise variiert wurden, haben im Herbst und Winter 2006 / 2007 ca. 1.500 Personen teilgenommen. Folgende, vom Konsortium nach Literaturrecherche (z.B. [6, 7, 8]) als komfort-relevant identifizierte Parameter wurden dabei untersucht: Luftdruck, Temperatur, relative Feuchtigkeit und Schall / Vibrationen sowie die Frischluftaustauschrate (weitere Details siehe Kapitel 2.4, Seite 12).

Die Reaktionen der Teilnehmer auf die dargebotenen Umweltbedingungen wurden dabei einerseits auf psychologischer Basis mittels Fragebögen, die zu verschiedenen Zeitpunkten (vor dem simulierten Flug, drei mal während und einmal nach dem Flug) ausgefüllt worden sind und andererseits auf physiologischer Basis erfasst. Dazu wurden – neben eines Bewegungsprofils per Aktometer (verantwortlich: Deutsche Forschungsanstalt für Luft- und Raumfahrt e.V. (DLR)) – die Reaktionen des Körpers (wie Herzrate, Hautpotential, Blutdruck, Puls sowie Sauerstoffgehalt des Blutes) kontinuierlich mit Hilfe von Datenloggern aufgezeichnet, die am Körper der Teilnehmer befestigt waren. Dieses geschah im Rahmen des ICE-Projekts unter der Leitung des Partners Medizinische Universität Wien, Österreich (MUV).

An vier der 35 simulierten Flüge nahmen zusätzlich medizinisch vorbelastete Passagiere (mit Herz-² bzw. Lungen-Problemen³) teil. Diese wurden (zusammen mit einer 'gesunden' Kontrollgruppe) vom Partner University College London, England (UCL) im Rahmen der Projektfragestellung auf z. B. mögliche Anzeichen einer potentiellen Thromboseentstehung hin mit Hilfe von Blutproben (speziell: Untersuchung auf *D-dimer* sowie *Interleukin-6*, Anzeichen für eine bevorstehende Blutverdickung als Indikator einer erhöhten Thrombosegefahr) und Lungenfunktionstests (*respiratory peak flow*) untersucht.

Ebenfalls zur gesundheitlichen Überwachung wurde im Rahmen dieser Dissertation bei 108 Teilnehmern jeweils vor und nach dem simulierten Flug die individuelle Ruhehörschwelle bestimmt. Dieses diente dazu, eine möglicherweise auftretende temporäre Hörschwellenverschiebung (TTS) zu dokumentieren. Eine detaillierte Beschreibung des Versuchs sowie die Auswertungen befinden sich in Kapitel 5 ab

 $^{^{2}}$ offizielle internationale Kodierung: NYHA grade 2 heart failure patients

 $^{^3 {\}rm offizielle}$ internationale Kodierung: COPD - MRC dyspnoea grade 2 patients

Seite 61.

Die resultierenden psychologischen und physiologischen Erkenntnisse wurden zur weiteren Auswertung mit den eingestellten Umgebungsbedingungen korrelliert, um statistisch gesicherte Aussagen treffen zu können, welche physikalischen Parameter unter welchen Expositionskombination das Wohlbefinden und die Gesundheit beeinflusst haben.

Die Ergebnisse wurden abschließend in Form eines Vorhersagemodells (PDM) vom Partner Stichting National Lucht en Ruimtevaartlaboratorium, Niederlande (NLR) zusammengefasst und dem Konsortium zur Verfügung gestellt. Damit steht im Rahmen der Ergebnisveröffentlichung des ICE-Projekts (*Dissemination*) verschiedenen Zielgruppen die Möglichkeit offen, entsprechend ihrer Belange die optimalen Komfortparameter für eine Flugzeugkabine (erstmalig auch unter Berücksichtigung kombinierter Effekte) vorhergesagt zu bekommen [5].

Eine weitere im Rahmen des Projekts vom Partner Avitronics Research, Griechenland (AVI) durchgeführte Vergleichsuntersuchung während zweier realer Mittelstreckenflüge (*in-flight tests*) zwischen Athen und Paris mit einem Airbus A320 der Fluglinie Hellas Airlines dienten der Evaluierung der Ergebnisse des PDM. Hierbei wurden die Umweltbedingungen an Bord eingeschränkt mit transportablen Messgeräten (für Temperatur, Druck, Feuchtigkeit und Schall) dokumentiert und eine Auswahl von Passagieren auf freiwilliger Basis mit Hilfe von Fragebögen befragt.

2.3 Verwendete Simulationsumgebungen

Untersuchungen des Einflusses von verschiedenen Umweltbedingungen auf das Komfortempfinden von Passagieren in einer Flugzeugkabine lassen sich verständlicherweise während realer Flüge nur sehr begrenzt durchführen, da besonders die Parametervariation in Reproduzierbarkeit und Dynamikbereich großen Einschränkungen unterliegt.

Um diese Fragestellung dennoch untersuchen zu können, verwendet man z. B. Simulationssysteme, die in der Lage sind, jederzeit die gewünschten Umweltbedingungen systematisch und reproduzierbar zu erzeugen. Systembedingt stellen Simulationen (bzw. Simulatoren) im Allgemeinen immer (nur) einen Teilausschnitt des Gesamtsystems bzw. der Realität dar, abhängig vom jeweiligen Grad der Abstraktion. Speziell bei Komfortuntersuchungen sind dabei besondere Anforderungen an das Simulationssystem zu beachten. Neben der technischen Realisierung zur definierten Darbietung der gewünschten Parameter ist die "optische" Darbietung ein wichtiger Bestandteil, um den Teilnehmer auf die Situation "an Bord" passend und umfangreich einzustimmen.

Im ICE-Projekt wurden folgende Kabinensimulationssysteme mit Schwerpunkt auf Komfortuntersuchungen in Flugzeugen verwendet: "Aircraft Cabin Environment (ACE)" (6 geplante Flüge) und "Flight Test Facility (FTF)" (29 geplante Flüge).

Aircraft Cabin Environment (ACE)

Der beim Building Research Establishment Ltd., England (BRE)⁴ im Rahmen der EU-geförderten Projekte HEACE ("Health Effects in Aircraft Cabin Environment", unter Leitung der Universität Oldenburg [8]) und FACE ("Friendly Aircraft Cabin", koordiniert von Alenia Aeronautica S.p.A., Italien [7]) entstandene Kabinensimulator ACE ermöglicht prinzipiell die Untersuchung der Auswirkung von Schall, Vibration, Temperatur, Luftfeuchtigkeit, Luftgeschwindigkeit und Luftaustauschrate auf den menschlichen Körper und dessen Wohlbefinden [9]. Er besteht aus einer Frontsektion eines realen Airbus A300-B4 mit einer Kapazität von ca. 40 Passagieren plus zwei Piloten und Kabinencrew. Der Simulator ist starr ohne Bewegungssystem aufgestellt. Abbildung 2.1 zeigt eine Funktionsskizze des Simulators aus dem Jahr 2004.

⁴siehe http://www.bre.co.uk, Juli 2009



Abbildung 2.1: ACE - Facility, Funktionsskizze aus dem Jahr 2004

Flight Test Facility (FTF)

Auf dem Gelände des Fraunhofer-Institut für Bauphysik (IBP) in Holzkirchen⁵ wurde 2006 ein Kabinensimulator auf Basis eines Airbus A310-200 starr in einer Unterdruckkammer aufgebaut. Es handelt sich dabei um eine 15.5 m lange reale Frontsektion (Cockpit und vorderer Bereich) eines *wide-body* Flugzeuges für bis zu 80 Passagiere mit einer 2-4-2-Bestuhlung. Die Fotos in Abbildung 2.2 zeigen links das Einsetzen des Rumpfs und rechts das fertig montierte System in der 30 m langen Druckkammer (Durchmesser: 9.6 m).

⁵siehe http://www.bauphysik.de/FTF, Juli 2009



Abbildung 2.2: Aufbau des Flight Test Facility (FTF) - Simulators (Quelle: www.bauphysik.de/FTF)

Dieser Simulator ermöglicht es, in einer nahezu realistischen Kabinenumgebung am Boden reproduzierbar und exakt einen Großteil der Umweltbedingungen einzustellen, die auf die Passagiere während eines realen Fluges in 10 km Höhe einwirken, insbesondere (Unter-) Druck. Diese können, entgegen den Rahmenbedingungen während eines regulären Fluges, nahezu beliebig und separat gegeneinander variiert werden, um den jeweiligen Einfluss auf den Menschen zu untersuchen.

Die Innenausstattung des verwendeten Flugzeugsegments wurde beibehalten, um eine auch optisch möglichst realistische Flugzeugumgebung zu gewährleisten (Innenansichten siehe Abbildung 2.3). Der Simulator bietet in der für die ICE-Versuche gewählten Konfiguration Platz für ca. 40 Passagiere (grau markierte Sitze in Abbildung 2.4) und zwei Flugbegleiter zur Betreuung der Passagiere im Rahmen der realitätsnahen Simulation. Im hinteren Teil bietet die Kabine Platz zur Unterbringung der Versuchsleiter und -beobachter. Die nicht markierten Sitze im vorderen Bereich sind mit Messgeräten belegt, die zur Überwachung der Umgebungsbedingungen dienen.

Zur Darbietung einer möglichst realistischen Sound- und Vibrationswiedergabe wurden durch die Universität Oldenburg in Kooperation mit dem An-Institut ITAP GmbH⁶ in jeder Reihe vier Lautsprecher verbaut, jeweils links und rechts im Seitenpanel oben bzw. unten. Das Schallfeld ist in der gesamten Kabine kalibriert auf sowohl gleichmäßiges Spektrum als auch auf einheitlichen Pegel (eine Abbildung des Spektrums befindet sich auf Seite 94, Kapitel 5.5).

⁶Institut für Technische und Angewandte Physik GmbH, http://www.itap.de, Juli 2009



Abbildung 2.3: Innenansichten des Flight Test Facility (FTF) - Simulators, links: Blick nach vorne, rechts: Blick nach hinten. Fotos: MUV, IBP

Zudem wurden durch die ITAP GmbH elektrodynamische Schwingerreger konzipiert und unter jedem Sitzplatz eingebaut, um jeden Sitz mit real gemessenen Sitzvibrationen anregen zu können. Die verwendeten (Original-) Flugzeugsitze können in erster Näherung als "single input - multiple output" System angesehen werden, d.h. die Anregung in nur einer Richtung produziert Vibrationen in alle drei Raumrichtungen auf der Kontaktstelle Sitz-Mensch.

Mit Hilfe einer Druckschleuse (nicht abgebildet) kann das Innere der Druckröhre (und damit die Simulationsumgebung) jederzeit betreten werden.



Abbildung 2.4: Bestuhlung des FTF-Simulators. Die grau markierten Plätze stehen den Teilnehmern der Untersuchung zur Verfügung (optimierte Konfiguration für das ICE-Projekt). Die verbleibenden Plätze sind entweder durch Messgeräte belegt (Zwischenplätze) oder für Versuchsleiter reserviert (hinterer Teil ab Reihe 7). Die zwei Zugänge zur Kabine befinden sich links in der Skizze. Die Abbildung wurde zur Verfügung gestellt vom Fraunhofer IBP.

2.4 Ausgewählte Umgebungsparameter und Variationsbereich

Folgende Umweltparameter und Variationsbereiche wurden vom Projekt-Konsortium mit Hilfe eines Versuchsdesignalgorithmus (weitere Details folgen in Kapitel 2.7, Seite 14) für die 29 Versuchsflüge im FTF identifiziert, um die jeweiligen und kombinierten Auswirkungen auf die Gesundheit und das Wohlbefinden der Passagiere zu untersuchen:

- äquivalente Kabinenhöhe⁷ (Druck): 0 ft⁸, 4'000 ft (875 hPa), 6'000 ft (810 hPa) und 8'000 ft (753 hPa)
- Temperatur: 21 °C, 23 °C und 25 °C
- relative Luftfeuchtigkeit: 10 %, 25 % und 40 %
- Schall⁹: 64 dB(A), 69 dB(A) und 74 dB(A)

⁷Der Luftdruck in der Passagierkabine wird in der Luftfahrt meist in einer äquivalenten Kabinenhöhe in Fuß (engl. feet, [ft]) angegeben. 0 ft entspricht dem Standardnormaldruck in Bodennähe (1013 hPa) und 8'000 ft ungefähr einem Druck von 753 hPa.

⁸"normaler" Umgebungsdruck in Holzkirchen entspricht $\approx 2'000$ ft (930 hPa).

⁹Verwendung eines realistischen Kabineninnengeräuschs bei unterschiedlichen Lautstärkepegeln

Dabei handelt sich einerseits um typische Werte während einer Langstreckenreise in einer Flugzeugkabine (8'000 ft, 23 °C, 10 % rel. Luftfeuchtigkeit und 74 dB(A)), andererseits aber auch um Werte, die einen anderen (evtl. höheren) Komfortlevel erwarten lassen. Von besonderem Interesse ist dabei der Einfluss des Kabinenluftdrucks (statt einer äquivalenten Kabinenhöhe von 8'000 ft (753 hPa) wird die Kabine nur auf 6'000 ft (810 hPa) "gehoben").

Im ACE wurden primär Untersuchungen mit unterschiedlichen Frischluftzufuhrraten durchgeführt. Die übrigen klimatischen Bedingungen wurden dabei systembedingt nicht konstant gehalten, sodass ein direkter Vergleich mit den Messungen im FTF nicht möglich ist. Die weitere Auswertung dieser Arbeit befasst sich daher schwerpunktmäßig nur mit den Ergebnissen aus den FTF-Simulationstests.

2.5 Nicht systematisch variierte Parameter

Neben den zuvor genannten Umweltbedingungen, die der systematischen Variation unterlagen, wurden weitere Parameter erfasst, die einen Einfluss auf das Wohlbefinden der Passagiere haben könnten, wie z.B. der Bereich "Luftqualität". Dazu zählen neben den Luftpartikeln ebenso die chemischen wie auch die biologischen Bestandteile der Luft, wie z. B. $CO, CO_2, O_3, VOCs^{10}, SVOCs^{11}, Kohlenwasserstoffe$ und Keime.

2.6 Messwertüberwachung und -speicherung

In der Kabine des FTF wurde eine große Anzahl von Messwertaufnehmern platziert, um die Umweltbedingungen an jedem einzelnen Sitz erfassen zu können. Dazu zählen hauptsächlich (Luft- und Strahlungs-) Temperatur sowie Luftfeuchigkeit, Luftströmung, Schall und Vibration. Andere Parameter wie Druck, Frischluftzufuhrrate (*ventilation rate*) und Luftaustauschrate (*recirculation rate*) wurden zentral an einer Stelle in der Kabine bzw. im Luftzulauf gemessen. Zusätzlich wurde der beim IBP entwickelte *DRESSMAN* [10] mit in die Kabine als "Passagier" in Reihe 6 gesetzt. Dieser verfügt an einigen Stellen über eine besondere Oberfläche, die in ihrer Reaktion auf Temperatur der menschlichen Haut nachempfunden wurde. Die Messdaten der per Ethernet miteinander verbundenen Messfühler wurden kontinuierlich

¹⁰volatile organic components - flüchtige organische Komponenten (VOC)

¹¹semi-volatile organic components - mittel-flüchtige organische Komponenten (SVOC)

außerhalb des Simulators in der Kontrollstation gespeichert. Die aufgenommenen Messwerte unterschieden sich dabei nur gering von den vorgesehenen und eingestellten Sollwerten [11].

Abbildung 2.5 zeigt einige Fotos [12] von ausgewählten Messfühlern und ihren Positionen.



Abbildung 2.5: Messfühler für Lufttemperatur, Strahlungstemperatur, Temperatur an der Innenverkleidung und Luftfeuchtigkeitssensor (von links oben nach unten rechts) [12].

2.7 Versuchsdesign

Grafisch in Abbildung 2.6 dargestellt, führen die Parameter- und Variationsvorgaben zu folgendem Versuchsdesign (jeder Punkt entspricht einem Messflug) [13]: Beginnend mit den drei möglichen Unterdruckbedingungen (in der Abbildung links dargestellt), ergeben sich neun Versuchsflüge, nimmt man weiter die drei Temperaturbedingungen hinzu (Mitte). Untersucht man nun auch die drei ausgewählten relativen Luftfeuchtigkeiten, so ergibt sich eine Messmatrix mit insgesamt 27 Versuchstagen (Bild rechts).



Abbildung 2.6: Voll faktorielles Versuchsdesign mit 3 Variablen und 3 verschiedenen Level (v.l.n.r.): Druck P (=3 Tests), Druck P + Temperatur T (=9 Tests), Druck P + Temperatur T + relative Luftfeuchtigkeit RH (=27 Tests) [13].

Wird zusätzlich der Einfluss unterschiedlicher Schallbedingungen als vierter Parameter mit untersucht, ergeben sich im voll faktoriellen Design 81 Versuchsflüge (siehe Abbildung 2.7).



Abbildung 2.7: Ein voll faktorielles Versuchsdesign mit insgesamt 4 Parametern (wie zuvor, zusätzlich mit 3 verschiedenen Lärmkonditionen) zu je 3 verschiedenen Level führt zu 81 Versuchsflügen [13].

Da dies weder zeitlich noch finanziell im Rahmen des ICE-Projekts realisierbar war, wurde stattdessen eine Variante des sog. *D-optimal design* [14] gewählt. Hierzu wird ein Design-Raum vordefiniert, bestimmt durch folgende Rahmenbedingungen:

- maximal mögliche Anzahl der Tests (technisch, finanziell, ...)
- minimal nötige Anzahl für eine valide Aussage des zu entwickelnden Vorhersagemodells

D-optimal Messdesign

D-optimal Messdesigns werden mit Hilfe von Software-Algorithmen erstellt und optimiert. Diese Verfahren finden dann Verwendung, wenn herkömmliche Designs aufgrund von Beschränkungen nicht angewendet werden können. Im Unterschied zu üblichen Standarddesigns (wie z. B. faktorielle oder teil-faktorielle Designs) sind D-optimal Designs nicht vollständig mit Messungen besetzt [15, Kapitel 5.5.2.1].

Durch die Festlegung auf maximal 28 Tests (+ 1 Reservetag) ergab sich das "modifizierte" D-optimal Messdesign, dargestellt in Abbildung 2.8.

Dabei wird als Optimierung nicht jede Kombination getestet, sondern nur "Schlüsselkombinationen". Zusätzlich zu den bisher definierten Versuchen wurden Vergleichstests bei Standardluftdruck (0 ft) hinzugenommen (unterste Ebene in Abbildung 2.8). Des Weiteren wurde die Schallkondition "69 dB(A)" nur ein Mal getestet. Die Flüge mit den Risikogruppen an Bord (rot markierte Messpunkte, gleichzeitig Symbol für die Messungen bei 64 dB(A)) fanden zudem nur bei bestimmten Umweltbedingungen statt (23 °C, 25 % bzw. 40 % rel. Feuchtigkeit, 74 dB(A) sowie 0 ft und 8′000 ft). Die Tabelle 2.1 auf Seite 23 zeigt die jeweiligen Konditionskombinationen für die einzelnen Messtage im Detail.



Abbildung 2.8: Finales Versuchsdesign im FTF. Im "modifizierten" D-optimal Messdesign, welches die gegebenen Randbedingungen im ICE-Projekt berücksichtigt, werden nur die blau markierten Kombinationen (bei 74 dB(A)) untersucht. Die gelbe Markierung zeigt die zusätzlich durchgeführte Untersuchung bei 69 dB(A), die rote Markierung zum einen die zusätzlichen Messungen bei 64 dB(A) sowie zum anderen die Messungen mit Personen aus den identifizierten Risikogruppen (diese jedoch ausschließlich bei 74 dB(A)). Mit der untersten Ebene wurde eine weitere Messkonfiguration bei normalem Umgebungsdruck (*ambient* \approx 930 hPa) eingeführt [13].

2.8 Teilnehmer

An den simulierten Flügen im FTF haben täglich wechselnde gesunde¹² Männer und Frauen aus der Region München teilgenommen. Die insgesamt 1.160 Probanden zwischen 18 und 93 Jahren wurden mit Hilfe einer örtlichen Personalagentur entsprechend der Projekt-Vorgaben ausgesucht (Altersverteilung siehe Abbildung 2.9). Für die Dauer des Versuchstages wurde den Teilnehmern eine Aufwandsentschädigung gezahlt.

An vier Messtagen (siehe Tabelle 2.1, Seite 23, risk 1 bis risk 4) fanden abweichend

¹²'gesund' im Sinne von keinen bekannten Herz- und Lungenproblemen, siehe folgender Absatz



Abbildung 2.9: Altersverteilung der Probanden im ICE-Versuch

Flüge statt, in denen die eine Hälfte der Passagiere 'gesund' war, der andere Teil je zur Hälfte unter Herzproblemen (*cardiovascular perturbation*) bzw. unter Lungenproblemen (*respiratory tract perturbation*) litt. Der Gesundheitszustand sowie die Flugtauglichkeit dieser Risikogruppen wurden im Vorfeld durch medizinisches Personal (gestellt von den Projektpartnern Civil Aviation Authority, England (CAA), DLR und UCL¹³) überprüft und getestet. Während dieser Messtage wurden von den Passagieren jeweils vor und nach dem simulierten Flug durch den Partner UCL Blutproben genommen, um ggfs. auftretende gesundheitliche Probleme, hervorgerufen durch die eingestellten Bedingungen an Bord, dokumentieren zu können. Im Falle von unvorhergesehenen medizinischen Komplikationen war eine Erstversorgung vor Ort gewährleistet.

2.9 Fragebögen

Um die Auswirkungen der Expositionen auf die Teilnehmer auch auf psychologischer Ebene zu erfassen, wurden zu verschiedenen Zeitpunkten umfangreiche Fragebögen, realisiert auf einem kompakten tragbaren Computer (Personal Digital

¹³Eine Liste der teilnehmenden Partner im ICE-Projekt befindet sich in Kapitel 4.8, Seite 58.

Assistant (PDA)), verteilt. Diese wurden insgesamt fünfmal während eines Testtages ausgehändigt: Zunächst vor dem Flug (generelle persönliche Daten), dreimal während der Testbedingung (Fragen zum aktuellen Empfinden) sowie am Ende des Fluges, in der die Erholungsphase stattfand (Abbildung 2.10 auf Seite 20 verdeutlicht den Ablauf schematisch).

Neben allgemeinen Fragen zu Alter, Geschlecht, Flugerfahrungen (unter Verwendung von Nominalskalen) und gesundheitlicher Selbsteinschätzung (Ordinalskalen) wurden auch die momentane Beurteilung der derzeit dargebotenen Umweltbedingung mit Hilfe von Fragen zur Selbsteinstufung ('Rating') abgefragt. Dazu wurden verschiedene Fragen zu den einzelnen Schwerpunkten der Untersuchung gestellt, wie z.B. zur Temperaturwahrnehmung, zur Feuchtigkeitsempfindung, zur Schallwahrnehmung sowie zur Wahrnehmung unterschiedlicher körperlicher (Krankheits-) Symptome, im Detail:

- allgemeiner Gesundheitsstatus (36 Fragen)
- allgemeines Wohlbefinden (27 Fragen)
- Sensitivität bzgl. Umweltbedingungen in der Selbsteinschätzung (jeweils einmal zu Anfang) (17)
- derzeitiges Wohlbefinden (mehrfach abgefragt) (27)
- derzeitige Einschätzung der Umweltbedingungen (14)
- derzeitige Beurteilung der Umweltbedingungen (20)
- derzeitige Beurteilung des allgemeinen Komfortempfindens (5)
- derzeit wahrgenommene Gesundheitsbeeinträchtigungen ("Symptome") (38)

Damit umfasste der Fragebogen insgesamt bis zu 183 Fragen.

Zur Beantwortung wurden primär vier- oder siebenstufige Ratingskalen verwendet: 4er Skalen, z. B. bei Fragen zur Gesundheit ("Ich leide unter Kopfschmerzen:" gar nicht (Antwort 1) bis sehr (Antwort 4)), bei Fragen zur Wahrnehmung der Umgebung ("Derzeit gibt es störende Geräusche von anderen Passagieren:" keine (1) bis viele (4)) oder Fragen zum Komfortempfinden ("Für mich ist die Sitzbreite … " unkomfortabel (1) bis komfortabel (4)).

7er Skalen, z. B. bei Fragen zur Komfort-Beurteilung der Umgebung ("Für mich ist die Temperatur am Kopf derzeit … " *zu kühl* (1) bis *zu warm* (7)) oder Fragen zum derzeitigen Wohlbefinden ("Derzeitig fühle ich mich schläfrig:" *überhaupt nicht* (1) bis *sehr stark* (7)).

Die Fragebögen wurden durch den Partner MUV erstellt [16], verteilt, ausgelesen und nach einer Vorverarbeitung (Details dazu werden in Kapitel 3.1 weiter vertieft) zur weiteren Auswertung zur Verfügung gestellt.

2.10 Messprogramm

Basierend auf den Ergebnissen der Versuchsplanung, Kapitel 2.4, Seite 12, wurde ein Ablaufplan gemäß Tabelle 2.1, Seite 23, für die Messflüge im FTF zwischen dem 6. November 2006 (Flug 1) und dem 26. Januar 2007 (Flug 29) festgelegt. Der Versuch vom 18. Januar 2007 (Flug 24) musste wegen einer Unwetterwarnung (Orkantief Kyrill) vorzeitig beendet werden. Dieser Messtermin wurde am 25. Januar als Flug 28 wiederholt (in der Tabelle jeweils mit (*) markiert).

2.11 Tagesablauf

Jeder Messtag mit einer Dauer von knapp 12 Stunden war – wie in Abbildung 2.10 zu sehen – in folgende fünf Sequenzen aufgeteilt [13]:



Abbildung 2.10: Ablaufplan der vorgestellten Messkampagne mit fünf Hauptsequenzen, rote Symbole (pre- und post-processing): medizinische Tests vor und nach dem simulierten Flug, schwarze Symbole (baseline, flight und post-baseline): 5× Ausgabe von Fragebögen [13].

pre-processing (2 h 15 min)

Zu Beginn ab ca. 8.15 Uhr wurde der Simulator im Rahmen der Vorbereitungen auf den jeweiligen Messtag technisch überprüft und für die geplanten Messungen eingerichtet. Die Teilnehmer wurden gegen 8.30 Uhr erwartet, über den Tagesablauf generell informiert (jedoch nicht über die geplanten Versuchsbedingungen) und ihr Einverständnis zur freiwilligen Teilnahme am Versuch eingeholt. Im Anschluss daran wurden die Teilnehmer von Mitarbeitern der Medizinischen Universität Wien sowie vom DLR mit den jeweiligen physiologischen Aufnahmeinheiten versehen. Während der Testtage mit den Risikogruppen wurden in dieser Phase auch die Blutproben und die Lungenfunktionstests (durch den Partner UCL) durchgeführt. Mit dem gemeinsamen Betreten des Simulators gegen 10.30 Uhr war die sog. pre-processing Phase beendet.

baseline (30 min)

An Bord des Simulators wurde zunächst eine *base-line* Phase eingehalten. Während dieser Phase wurde der erste (generelle) Fragebogen an die Passagiere zur Beantwortung ausgeteilt. Diese 30 Minuten wurden zudem zur Messung der physiologischen "Ruhebedingung" genutzt.

flight (7 h)

Die anschließende *flight* Phase begann mit der akustischen Darbietung einer original aufgenommen Startsequenz eines Airbus A320-Flugzeuges, mit *taxiing* (10 min), *take-off* (7 min) und Steigflugphase (*climb*, 13 min). Nach diesen 30 Minuten wurde der Schallpegel gegen 11.30 Uhr mit Erreichen der 6 h langen *cruise* Phase auf den vorgesehenen Tageswert eingestellt (64 dB(A), 69 dB(A) oder 74 dB(A)). Ebenso hat das Kabinenklima (Luftfeuchtigkeit, Temperatur sowie der Luftdruck an Bord¹⁴) den jeweiligen Tagestestwert erreicht. Während dieser Phase wurden die physiologischen Daten der Passagiere kontinuierlich aufgezeichnet. Ebenso wurden die psychologischen Daten dreimal mit Hilfe der per PDA realisierten Fragebögen erhoben. In der abschließenden 30-minütigen Anflug- und Landephase wurde eine original Landesequenz einer Airbus A320-Maschine akustisch eingespielt, darunter 10 min *descent*, 10 min *touch-down* und 10 min *taxiing* zum Terminal.

¹⁴An den ersten zwei Messtagen wurde abweichend ein variables Druckprofil getestet (von 4'000 ft zu Beginn über 6'000 ft zu 8'000 ft zum Ende hin bzw. umgekehrt.

post-baseline (30 min)

In der Zeit von 18.00 Uhr bis 18.30 Uhr wurde im Rahmen der Erholungsphase unter Ruhebedingungen die Regenerierung der Teilnehmer sowohl physiologisch gemessen als auch mit Hilfe eines abschließenden Fragebogens ermittelt.

post-processing (1 h 30 min)

In der letzen Phase fand das Deboarding und das Abkabeln der Passagiere statt (sowie eine Blutentnahme und Lungenfunktionstests während der Tests mit den Risikogruppen). Mit der Verabschiedung der Probanden gegen 19 Uhr und Abschluss der Datensicherung gegen 20 Uhr war der Messtag beendet.
Tag	Versuch	Druck Feuchtigk		Temperatur	\mathbf{SPL}
		ft % rel. F.		$^{\circ}\mathrm{C}$	dB(A)
1	Druck Profil 1	_	10	23	74
2	Druck Profil 2	—	10	23	74
3	0 ft-control 1	0	10	23	74
4	0 ft-control 2	0	25	23	74
5	0 ft-control 3	0	40	23	74
6	D-optimal 6	6000	10	23	74
7	69 dB improvement	8000	10	23	69
8	D-optimal 7	6000	25	23	74
9	D-optimal 4	8000	25	23	74
10	D-optimal 5	8000	40	23	74
11	D-optimal 13	4000	10	21	74
12	D-optimal 8	4000	10	23	74
13	D-optimal 11	4000	10	25	74
14	D-optimal 12	4000	40	21	74
15	D-optimal 10	4000	40	25	74
16	D-optimal 1	8000	10	21	74
17	D-optimal 2	8000	10	23	74
18	D-optimal 3	8000	10	25	74
19	D-optimal 14	8000	40	21	74
20	D-optimal 9	8000	40	25	74
21	64 dB-control 4	8000	40	23	64
22	64 dB-control 3	8000	10	23	64
23	64 dB-control 1	0	10	23	64
24	64 dB-control 2 (*)	0	40	23	64
25	risk 4	8000	40	23	74
26	risk 3	8000	10	23	74
27	risk 1	0	10	23	74
28	reserve (64 dB-c2) (*)	0	40	23	64
29	risk 2	0	40	23	74

Tabelle 2.1: Messplan für die FTF-Tests (Spalten 1 und 2) sowie beschlossene Sollwerte für Druck, rel. Luftfeuchtigkeit, Temperatur und Schall (Spalten 3-6). Flug 24 wurde am Reservetermin 28 wiederholt (jeweils markiert (*)).

2.12 Datenerhebung und -speicherung

Sämtliche Messdaten, die im Rahmen der Versuche entstanden sind, wurden auf Beschluss der teilnehmenden Projektpartner anonymisiert in einer Datenbank ähnlichen Struktur gesichert (Microsoft Excel 2003). Dazu wurde zeilenweise jedem Probanden (ein-eindeutig identifizierbar über eine Identitätskennung) sämtliche gemessenen Werte der Umgebungsparameter sowie die Fragebogenantworten zugeordnet. Aufgrund der Spaltenbeschränkungen in der verwendeten MS-Excel-Version wurden dazu verschiedene Tabellen erzeugt, die jeweils einen Untersuchungsschwerpunkt beinhalteten (z.B. physiologische Daten, Luftqualitätsdaten, Fragebogendaten, etc.).

Nach eingehender Überprüfung sämtlicher Daten auf Vollständigkeit, Konsistenz und Aussagekraft durch die zuständigen Partner wurde eine Kopie des gesamten Datensatzes erstellt und in diesem mit der Datenreduktion begonnen, um z. B. redundante Einträge für die weiterführende Datenanalyse zu entfernen. Dazu gehörten u. a. auch chemische Luftqualitätsparameter, die zwar überwacht wurden, jedoch über die Zeit der Versuche keine Veränderung aufwiesen. Ebenso wurden die Messparameter aus der Arbeitskopie entfernt, bei denen die gemessenen Werte dauerhaft unterhalb der für den Menschen wahrnehmbaren Grenzwerte lagen. Des Weiteren wurden die Fragebogendaten mit Hilfe verschiedener statistischer Methoden wie der Clusteranalyse hinsichtlich einer möglichen Zusammenfassung analysiert.

Abschließend wurden die so erstellten und reduzierten Einzeltabellen vom verantwortlichen Partner NLR zu einer Gesamttabelle zusammengefasst, die im Folgenden allen Partnern für weitere Analyse- und Modellierungszwecke zur Verfügung stand.

Kapitel 3 beschreibt im Anschluss die im Rahmen dieser Arbeit durchgeführten Auswertungsschritte.

Basierend auf den jeweiligen Ergebnissen der Auswertungen der involvierten Partner wurde von NLR abschließend das PDM entwickelt, welches als Vorhersagemodell für den Komfort in Flugzeugkabinen das Ergebnis des ICE-Projekts darstellt.

3 ICE II – Auswertung

Im folgenden Kapitel wird ein grundlegender Teil der durchgeführten Auswertungen der zur Verfügung gestellten Fragebogendaten beschrieben. Schwerpunkt des Kapitels bildet die Beschreibung eines hierfür entwickelten Umskalierungsverfahrens, welches notwendig ist, um die gelieferten Fragebogendaten für eine weitergehende Analyse und zur anschließenden Parametrisierung hinsichtlich der Modellierung vorzubereiten. Bei diesem Umskalierungsverfahren wird die Fragebogenantwort auf einer n-kategorialen Quellskala anteilig einer neuen m-kategorialen Zielskala zugeordnet. Die Beschreibung der nachfolgenden Parametrisierung folgt dann ausführlich in Kapitel 4.

3.1 Vorbemerkung

Die vorliegenden Fragebogendaten wurden hinsichtlich einer späteren Modellbildung (PDM) mittels statistischer Methoden (wie Hauptkomponentenanalyse - engl. PCA) sowie durch Mittelung von Antworten mit ähnlichem Fragekomplex vorausgewertet und vom Partner MUV¹ zur Verfügung gestellt.

Die gelieferten Daten weisen aufgrund dieser Vorverarbeitung teilweise Skalierungen auf, die sich von den Ursprungsskalen (z. B. 4-stufige Kategorialskala) und ihrer Bedeutung unterscheiden. Abbildung 3.5, Seite 30, zeigt ein Beispiel einer umskalierten Antwortverteilung auf einer Skala von '0' bis '1'.

Um die Antwortverteilungen mittels einer standardisierten Parametrisierung in das zu entwickelnde Vorhersagemodell eingliedern zu können, werden diese Daten dazu wieder die Ursprungsform transformiert bzw. umskaliert. Der Prozess der Parametrisierung wird nachfolgend in Kapitel 4 ausführlich beschrieben.

Im Folgenden werden zunächst grundsätzliche Überlegungen dargestellt, gefolgt von der Definition sowie der Anwendung der entwickelten Zuordnungsvorschrift auf die zu analysierenden Fragebogendaten.

¹Medizinische Universität Wien, Österreich

3.2 Psychologische Analyse mittels Fragebögen

Wie bereits in Kapitel 2.9 beschrieben, wurden zu verschiedenen Zeiten während des simulierten Fluges Fragebögen an die Teilnehmer ausgehändigt. Für die weitere detaillierte Auswertung innerhalb des ICE-Projekts waren dabei folgende Vorüberlegungen notwendig:

Ausgehend von der Beurteilung einer bestimmten Situation auf einer n-stufigen Ratingskala mit Rating-Index k von 1 bin n, erhält man von einem Teilnehmer t auf die Frage x_k eine Antwort entsprechend der Einstufung auf der Skala (z. B. durch Ankreuzen). Trifft man nun die Aussage über die Häufigkeit N, mit der eine Ratingantwort gegeben wurde, so lässt sich dies darstellen mit

$$N_t(x_k(t)) = \begin{cases} 1 & \text{, für } k = \text{ beurteilter Rating-Index} \\ 0 & \text{, sonst} \end{cases}$$
(3.1)

Betrachtet man nicht nur die Antwort eines einzelnen Teilnehmers, sondern die Antworten aller an einem (Tages-) Versuch teilnehmenden Personen (T) auf eine gestellte Frage $(x_k(T))$, so ergibt sich eine entsprechende Häufigkeitsverteilung N_T . Dazu ein Beispiel:

Es wird die Frage gestellt "Wie zufrieden sind sie?" mit den Antwortmöglichkeiten von 1 ('gar nicht') bis 4 ('sehr'). Abbildung 3.1 zeigt die Antworten von drei verschiedenen Teilnehmern. Teilnehmer 1 (t = 1) kreuzt, wie Teilnehmer 2 auch, auf der gegebenen (n =) 4-stufigen Skala Antwort '4' an. Für die gestellte Frage x_k ergibt sich somit für Teilnehmer 1 $x_4(t = 1) = 1$ (sowie $x_{1..3}(t = 1) = 0$) und für Teilnehmer 2 $x_4(t = 2) = 1$ sowie ebenfalls $x_{1..3}(t = 2) = 0$. Teilnehmer 3 antwortet hingegen mit '1': Damit ergibt sich $x_1(t = 3) = 1$ (sowie $x_{2..4}(t = 3) = 0$). Aufsummiert ergäbe dies eine Häufigkeitsverteilung von $1 \times x_1$, $0 \times x_2$, $0 \times x_3$ und $2 \times x_4$.

_	k =	12	34	
t = 1	$\square \rangle$		X	x ₁₃ = 0; x ₄ = 1
t = 2			X	x ₁₃ = 0; x ₄ = 1
t = 3		X		$x_1 = 1; x_{2.4} = 0$

Abbildung 3.1: erstes Beispiel für ein mögliches Antwortverhalten: 3 Einzelbeispiele für $x_k(t)$ mit t = 1 ... 3

In Abbildung 3.2 ist als zweites Beispiel die Häufigkeitsverteilung N der Antworten von nun 15 Teilnehmern auf eine Beispielfrage $(N_{T=15}(x_k(T)))$ als Histogramm dargestellt.



Abbildung 3.2: Beispiel für ein mögliches Antwortverhalten: Häufigkeitsverteilung der Antworten $x_k(T)$ (rechts dargestellt) von 15 Teilnehmern (links und Mitte) auf die Frage "Wie zufrieden sind Sie?".

Aus Abbildung 3.2 ergibt sich somit folgende Antwortverteilung der 15 Teilnehmer:

$$N_{15}(x_1) = 1$$

$$N_{15}(x_2) = 2$$

$$N_{15}(x_3) = 3$$

$$N_{15}(x_4) = 9$$

3.3 Ratingskalen

In den im ICE-Projekt verwendeten Fragebögen wurden teilweise unterschiedliche Ratingskalen für gleiche oder ähnliche Fragestellungen verwendet (z. B. Bewertung der Temperatur in der Kabine), um störende Nebeneinflüsse wie Ablenkungseffekte bzw. Reihenfolgeeffekte zu minimieren. Ebenso werden Wiederholungs- oder Kontrollfragen verwendet.

Zur Filterung und zur Datenreduktion wurden zwei unterschiedliche Varianten angewendet (vgl. Kapitel 3.1):

1. Durch die Verwendung statistischer Methoden wie z. B. der Hauptkomponentenanalyse (PCA) werden Antworten gleichgerichteter Fragenkomplexe zusammengefasst. Als Ergebnis stehen sogenannte **Faktoren** zur Verfügung.

2. Ein anderer Schritt zur Datenreduktion ist die Zusammenfassung von Antworten einzelner Fragen zu sogenannten **Indizes**, basierend auf einer 'semantischen' Grundlage (durchgeführt vom Partner MUV). Dabei werden nach vorheriger Normierung die Antworten auf Fragen per Mittelwertbildung vereint, die gleiche Themenkomplexe behandeln, also basierend auf einem inhaltlichen Schwerpunkt der Fragen.

3.4 Normierung

An dieser Stelle sei auf ein Problem hingewiesen, welches bei der Verarbeitung von Ordinalskalen bestehen kann [17]. Bei der Ordinalskala, wie sie bei Ratingfragen verwendet werden, wird die Merkmalsausprägung genau einer Kategorie zugeordnet, die mit einer Bezeichnung (komfortabel ... unkomfortabel) oder einer zu ihrer Repräsentation verwendeten Zahl ('1' ... '4') beschriftet werden kann. Ein weiterer Schritt ist das Definieren einer Rangfolge (komfortabel ist besser als weniger komfortabel und besser als unkomfortabel). Statistische Analysen von Ordinaldaten können sich daher auch immer nur auf Auswertungen von Ranginformationen beziehen. Anders ist dies bei der Intervallskala, wo ein empirisches Relativ erforderlich ist. Dies ermöglicht, Abstände zwischen Antwortkategorien anzugeben und damit auszuwerten. So werden gleich große Merkmalsunterschiede durch äquidistante Zahlen abgebildet, was die Berechnung von Mittelwerten und anderen statistischen Maßnahmen ermöglicht. Bortz [17, Seite 70] weisst in diesem Zusammenhang auf die "übliche Forschungspraxis [hin, nach der] auf eine empirische Überprüfung der jeweiligen Skalenaxiomatik verzichtet wird". Vielmehr wird *durch Vertrauen* davon ausgegangen, das jeweilige untersuchte Merkmal ließe sich auf einer Intervallskala messen. Damit werden erheblich differenziertere statistische Auswertungen ermöglicht.

So wird auch in dieser Untersuchung angenommen, die untersuchten Merkmale lassen sich per Intervallskala messen, so dass eine Mittelwertbildung und -auswertung im weiteren Verlauf anwendbar wird.

Werden nun im Rahmen dieser Datenreduktion einpolige Skalen mit zweipoligen (oder bipolare) Skalen kombiniert, bedarf es einer vorherigen Transformation. Fragt man z. B. nach der Zufriedenheit mit der Temperatur und bietet eine einpolige 4-er Skala an (Abbildung 3.3, oben), so entspricht der Antwort a = 1 hier der Beurteilung 'unkomfortabel', während die Antwort a = 4 der Beurteilung 'komfortabel' entspricht. Bei einer bipolaren 7er-Skala (Abbildung 3.3, unten) ergeben sich z. B. die Antworten b = 1 für 'zu kalt', b = 4 für 'komfortabel' und b = 7 für 'zu warm'.

Um nun einen kombinierten (und auf 1 normierten) Temperatur-Index I_{temp} zu bilden, der diese beiden Fragen zusammenfasst (auf einer Skala von 0 bis 1 mit '0':



3 ICE II – Auswertung

Abbildung 3.3: Beispiel für ein mögliches Antwortverhalten, oben: 4er Skala von 'unkomfortabel' (1) bis 'komfortabel' (4), unten: 7er Skala von 'zu kalt' (1) bis 'zu warm' (7). Es wurde auf beiden Skalen eine optimale Zufriedenheit angekreuzt.

unkomfortabel (zu kalt oder zu warm), '1': komfortabel, weder zu warm noch zu kalt), ist beispielsweise folgende Umrechnung der Skalen möglich:

$$\tilde{a} = \frac{3 - |a - 4|}{3}$$
 für einpolige Skalen bzw. (3.2)

$$\tilde{b} = \frac{3 - |b - 4|}{3}$$
 für zweipolige Skalen (3.3)

Die sich daraus ergebende Umskalierung zeigt Abbildung 3.4.



Abbildung 3.4: wie Abbildung 3.3, jedoch normiert und umskaliert von einem Wertebereich '1' bis '4' bzw. '1' bis '7' auf einen Wertebereich von '0' bis '1'. Das Anwortverhalten entspricht wieder einer optimalen Zufriedenheit mit der Temperatur.

Dabei ist zu beachten, dass die Zuordnung gemäß Gleichung (3.3) keine eineindeutige Rücktransformation in eine 7-Kategorienskala mehr zulässt. Dies wird aber bei diesen kombinierten Indizes im Rahmen der Projektauswertung und zur Implementierung in der Vorhersagesoftware in Kauf genommen. I_{temp} wird abschließend über die Berechnung des Mittelwert gebildet:

$$I_{temp} = Mittelwert(\tilde{a}, \tilde{b}) \tag{3.4}$$

Bei der Zusammenfassung zweier oder mehr solcher Fragen mittels Mittelwertbildung zu einem kombinierten Index können somit auch eine Anzahl von weiteren Zwischenwerten entstehen, wie in Abbildung 3.5 zu sehen ist. Sie zeigt anhand eines weiteren Beispiels die Anwortverteilung von 40 Befragten auf eine Frage. Für 20 der 40 Teilnehmern zeigt sich eine Präferenz hin zur Antwort '1', für 10 Teilnehmern ergibt sich '0.83', welche sich rechnerisch aufgrund der Mittelwertbildung aus den ursprünglichen Antworten '0.67' und '1' berechnet. Die verbleibenden 10 Teilnehmer verteilen sich auf die Antworten '0.5', '0.33' und '0.17'.



Abbildung 3.5: Beispiel einer Häufigkeitsverteilung mit 40 Teilnehmern, normiert und umskaliert auf den Antwortbereich '0' bis '1'.

Ein Großteil der im Projekt gelieferten und zur Analyse zur Verfügung gestellten Fragebogendaten wurde nach diesem Verfahren durch den Partner MUV vorverarbeitet.

3.5 Rücktransformation

Für die weitere Auswertung wird nun die Häufigkeitsverteilung, welche die Häufigkeiten für die generisch erzeugten (intervallskalierten) Antworten 0 bis 1 angibt (siehe Abbildung 3.5), wieder abgebildet auf die ursprüngliche (ordinale) 4er Antwort-Skala mit ihrer jeweiligen Zuordnung / Bedeutung ('1': unkomfortabel, '4': komfortabel, wie in Abbildung 3.3 oben).

Dies ist insbesondere dann hilfreich, wenn nach Befragung einer Teilnehmergruppe die Häufigkeitsverteilung eines solchen nachträglich berechneten Index analysiert (und später zur Modellierung parametrisiert) werden soll.

Diese Rückabbildung der Antworten erfolgt mit Hilfe einer Zuordnungsvorschrift ϕ (Gleichung (3.5)), in der mit geeigneten Algorithmen die jeweilige (auf 1 normierte) Antwort aus dem Skalenbereich \tilde{X} anteilig einer Antwort(-kategorie) aus dem ursprünglichen 4er Skalensystem \tilde{Y} (z. B. durch Diskretisierung) zugeordnet wird.

$$\phi = \tilde{X} \to \tilde{Y} \tag{3.5}$$

Abbildung 3.6 skizziert den zuvor beschriebenen Ablauf (Normierung nach Gleichung (3.2) bzw. (3.3), Erstellung eines kombinierten Index per Mittelwert (Gleichung (3.4)) und abschließender Rücktransformation ϕ , Gleichung (3.5)).

3.5.1 Skalenproblematik (Rück-)Transformation

Eine Zuordnung von einer Intervallskala hin zu einer Ordinalskala wird oftmals aufgrund der unterschiedlichen Skalenniveaus als "unzulässige Skalentransformation" betrachtet, wobei es nach Rohwer [18, Seite 134] "keinen erkennbaren Gesichtspunkt [gibt], um an dieser Stelle 'zulässige' von 'unzulässigen' Skalentransformationen zu unterscheiden". Im Folgenden wird der Ansatz verfolgt, die intervallskalierten Rohdaten (mit n Kategorien, $n \in \mathbb{N}$) mittels Neueinteilung in eine diskretisierte Intervallskala (mit m Kategorien, $m \in \mathbb{N}$) umzuwandeln.

Eine Zuordnung mittels Diskretisierung durch Anwenden der Gaußklammern² auf die Antworten in der Ursprungsskala \tilde{X} mit $\lfloor x_k \rfloor$ oder $\lceil x_k \rceil$ erweist sich dabei jedoch ebenso ungeeignet wie das (kaufmännische) Auf- bzw. Abrunden auf die nächste ganze Zahl (Kategorie).

Stattdessen wird im Folgenden eine Zuordnung ϕ definiert, bei der die Antwort auf der Quellskala anteilig auf die Zielskala abgebildet wird.

²auch Abrundungs- | | bzw. Aufrundungsfunktion [] genannt



Abbildung 3.6: Ablaufskizze zur Verarbeitung der Fragebogendaten am Beispiel zweier Skalen: (a) Normierung nach Gleichung (3.2) bzw. (3.3), (b) Erstellung eines kombinierten Index per Mittelwert (Gleichung (3.4)) und (c) abschließender Rücktransformation mittels Zuordnungsvorschrift ϕ (Gleichung (3.5)). Der Mittelungsprozess kann sich auch über weitere Skalen erstrecken, wenn es für die jeweilige Indexerstellung notwendig ist (rechts in grau angedeutet). Dabei ist zu beachten, dass bei bipolaren Skalen nach einer Normierung keine eineindeutige Rückzuordnung mehr möglich ist.

3.6 Zuordnungsvorschrift ϕ

3.6.1 Quellskala

Im ursprünglichen (Fragebogen-) Bereich \tilde{X} gibt es pro Frage eine Antwort x aus dem möglichen Antwortbereich k mit $k = 1 \dots n$ ($k \in \mathbb{N} \setminus \{0\}$, n entspricht der Anzahl der maximal gültigen Kategorien) und die entsprechende Häufigkeit $N(x_k)$ der Antwort x_k , wobei die Häufigkeit $P(x_k)$ einer Antwort x_k zwischen 0 und der maximalen Anzahl von Probanden T liegen kann (klassische Häufigkeitsverteilung).

3.6.2 Zielskala

Die Antwortkategorie y_l im Bereich der Zielskala (\tilde{Y}) mit l = 1 ... m $(l \in \mathbb{N} \setminus \{0\})$ tritt mit einer Häufigkeit $M(y_l)$ auf. Auch hier kann $M(y_l)$ Werte zwischen 0 und T annehmen (entspricht ebenfalls einer klassischen Häufigkeitsverteilung).

Es ergibt sich daraus eine Anzahl von maximal möglichen Kategorien Y mit

$$Y = y_m - y_1 + 1 = m. (3.6)$$

Zusammen mit dem Startwert der Quellskala x_1 und dem maximal (laut Fragebogen bzw. nach Umskalierung) möglichen Endwert (x_n) lässt sich die Intervalllänge Int wie folgt bestimmen:

$$Int = \frac{x_n - x_1}{Y - 1}.$$
 (3.7)

3.6.3 Grenzen

Für die untere und obere Grenze des Intervalls (Gleichung (3.7)) um die ursprüngliche Antwort x_k auf der <u>Quellskala</u> ergibt sich unter Zuhilfenahme von Gleichung (3.7) folgender Zusammenhang:

Untere Grenze ($[\leftarrow)$:

$$x_k \stackrel{[\leftarrow}{=} = x_k - \frac{Int}{2} \tag{3.8}$$

Obere Grenze $(\rightarrow]$:

$$x_k \stackrel{\rightarrow]}{=} x_k + \frac{Int}{2}, \qquad (3.9)$$

grafisch in Abbildung 3.7 dargestellt.



Abbildung 3.7: Antwort x_k mit den Grenzen $x_k \stackrel{[\leftarrow}{}$ und $x_k \stackrel{\rightarrow}{\rightarrow}$

Für die <u>Zielskala</u> gelten entsprechend für die untere und obere Intervallgrenze um die Kategorie l folgende Zusammenhänge (grafisch in Abbildung 3.8 dargestellt):

$$y_l \stackrel{[\leftarrow}{=} x_1 - \frac{Int}{2} + (l-1) \cdot Int$$
 (3.10)

$$y_l \stackrel{\rightarrow}{}^{]} = x_1 - \frac{Int}{2} + l \cdot Int \tag{3.11}$$



Abbildung 3.8: Antwort y_l mit den Grenzen $y_l \stackrel{[\leftarrow]}{\leftarrow}$ und $y_l \stackrel{\rightarrow]}{\rightarrow}$

3.6.4 Zuordnung

Im Folgenden wird nun getestet, inwieweit eine (Fragebogen-)Antwort x_k mit ihren zugehörigen Intervallen ($x_k \ [\leftarrow < x_k < x_k \]$) in eine Zielkategorie y_l mit ihrem jeweiligen Intervallbereich ($y_l \ [\leftarrow < y_l < y_l \]$) überlagert. Dazu werden zwei getrennte Bereiche betrachtet, einmal das Verhalten an der unteren (Gleichung (3.12)) und einmal entsprechend an der oberen Grenze (Gleichung (3.13)).

Verhalten untere Grenze:

$$A_{[x_k,y_l]} = \begin{cases} \frac{y_l \stackrel{\rightarrow}{\longrightarrow} - x_k \stackrel{[\leftarrow}{\longrightarrow}}{Int} & \text{, falls } (x_k \stackrel{[\leftarrow}{\longrightarrow} \leq y_l \stackrel{\rightarrow}{\longrightarrow}) \text{ und } (x_k \stackrel{[\leftarrow}{\longrightarrow} > y_l \stackrel{[\leftarrow}{\longrightarrow}) \\ 0 & \text{, sonst} \end{cases}$$
(3.12)

Verhalten obere Grenze:

$$B_{[x_k,y_l]} = \begin{cases} \frac{x_k \rightarrow] - y_l }{Int} & \text{, falls } (x_k \rightarrow] \le y_l \rightarrow]) \text{ und } (x_k \rightarrow] > y_l (\leftarrow) \\ 0 & \text{, sonst} \end{cases}$$
(3.13)

Grafisch wird dieses in Abbildung 3.9 veranschaulicht.



Abbildung 3.9: Antwort x_k wird mit $A_{[x_k,y_{l-1}]}$ anteilig der Kategorie y_{l-1} und mit $B_{[x_k,y_l]}$ der Kategorie y_l zugeordnet.

Für Kategorie y_l ergibt sich nach (3.12) ein Anteil $A_{[x_k,y_l]}$ von 0 (Bedingung 2 nicht erfüllt) und nach (3.13) ein Anteil $B_{[x_k,y_l]}$ von $\frac{x_k \rightarrow l - y_l}{Int}$. Für Kategorie y_{l-1} ist $A_{[x_k,y_{l-1}]} = \frac{y_{l-1} \rightarrow l - x_k}{Int}$ und $B_{[x_k,y_{l-1}]} = 0$.

Der sich letztlich aus der Antwort x_k für die Kategorie y_l ergebende Anteil S_{x_k,y_l} ergibt sich aus der Summe der Anteile $A_{[x_k,y_l]}$ und $B_{[x_k,y_l]}$ der beiden Grenzbetrachtungen aus (3.12) und (3.13), multipliziert mit der Häufigkeit N der Antwort x_k wie folgt:

$$S_{[x_k,y_l]} = (A_{[x_k,y_l]} + B_{[x_k,y_l]}) \cdot N(x_k)$$
(3.14)

Summiert man nun alle Anteile $S_{[x_k,y_l]}$, die sich abbilden lassen auf die jeweilige Klasse der Zielskala y_l , erhält man die neu zugeordnete Häufigkeit $M(y_l)$:

$$M(y_l) = \sum_{k=1}^{n} S_{[x_k, y_l]}$$
(3.15)

3.7 Beispieldarstellung

Verwendet man die Beispielverteilung aus Abbildung 3.5 und rechnet diese mit Hilfe der Zuordnungsvorschrift ϕ in eine 4er Skala mit Kategorien von 1 bis 4 um, so ist bei einer möglichen Umsetzung mit einer Tabellenkalkulationssoftware eine Darstellung gemäß Abbildung 3.10 denkbar. Für die Umskalierung aller anfallenden Daten erfolgte eine Umsetzung der Zuordnungsvorschrift in MATLAB³

		Quellsk	ala					Ziels	kala			
		Quellak	ala					21013	Kulu			
					=	: 1	=	= 2	=	= 3		= 4
			untere	obere	Y _{I=1}	= 0	$y_{1=2} =$	0,333	y ₁₌₃ =	0,667	y 1=4	= 1
	Antwort	Häufigkeit	Grenze	Grenze	y ₁₌₁ [<=	y ₌₁ ^{=>]}	y ₁₌₂ [<=	y ₁₌₂ =>]	y ₁₌₃ [<=	y ₁₌₃ =>]	y ₁₌₄ [<=	y ₁₌₄ =>]
k	X _k	N(X _k)	X _k ^{[<=}	x _k ^{=>]}	-0,167	0,167	0,167	0,500	0,500	0,833	0,833	1,167
1	0,00	0	-0,167	0,167		B(x1,y1)=1,0						
2	0,17	1	0,000	0,333	A _(x2,y1) =0,5			B _(x2,y2) =0,5				
3	0,33	2	0,167	0,500				B _(x3,y2) =1,0				
4	0,50	2	0,333	0,667			A(x4,y2)=0,5			B(x4,y3)=0,5		
5	0,67	5	0,500	0,833					A(x5,y3)=1,0			
6	0,83	10	0,667	1,000					A(x6,y3)=0,5			B(x6,y4)=0,5
7	1,00	20	0,833	1,167								B _(x7,y4) =1,0
					M(y _{i=1}) = 0,5	M(y _{i=2}) = 3,5	M(y ₁₌₃)	= 11,0	M(yı	.4) = 25

Abbildung 3.10: Screenshot zur Umsetzung der Zuordnungstabelle in einer Tabellenkalkulationssoftware (mit zusätzlichen Textergänzungen): Für eine Antwortverteilung auf einer 7-stufigen Quellskala von '0' bis '1', links im Bild, werden zunächst die unteren und oberen Grenzen (4. und 5. Spalte) für jede Antwort x_k (Spalte 2) berechnet. Im rechten 4-stufigen Zielskalenbereich wird, entsprechend der anteiligen Zuordnung zur jeweiligen neuen Kategorie, der zugehörige Anteil ausgewiesen. Die Multiplikation mit der beobachteten Häufigkeit $N(x_k)$ im Bereich der Quellskala führt abschließend zur umskalierten Häufigkeit $M(y_l)$ für den Bereich der Zielskala.

Abbildung 3.11 zeigt links den Verteilungsplot der ursprünglichen Antwortverteilung auf der normierten Quellskala '0' bis '1' und auf der rechten Seite die umskalierte Verteilung auf die Kategorien '1' bis '4'.

³Version 7.5.0.342 (R2007b), The MathWorks



Abbildung 3.11: Ergebnis der Umskalierung. Links: ursprüngliche Antwortverteilung auf der normierten Quellskala '0' bis '1'. Rechts: auf die Kategorien '1' bis '4' umskalierte Verteilung.

3.8 Zusammenfassung und weiteres Vorgehen

In diesem Kapitel wurde eine Methode vorgestellt, die es ermöglicht, aus Rohdaten berechnete, kombinierte Indizes mit einer beliebigen (Intervall-) Skala geeignet auf eine neue diskretisierte Intervallskala (und damit wieder zurück zu einer Art Ordinalskala) zuzuordnen. Diese Umskalierung von Fragebogendaten hinsichtlich zu diskreten Verteilungsfunktionen dient als notwendige Grundlage für die weitere Auswertung hinsichtlich der im folgenden Kapitel 4 beschriebenen Methode, die beobachteten Antwortverteilungen bestmöglich zu parametrisieren.

4 ICE III – Distribution of Subjective Assessments in a Controlled Aircraft Environment

Nach Vorstellung der durchgeführten Versuche im ICE-Projekt (Kapitel 2) und der Datenaufbereitung hinsichtlich einer Diskretisierung der zur Verfügung gestellten Antwortverteilungen in Kapitel 3 folgt nun im vorliegenden 4. Kapitel die Beschreibung der zur Parametrisierung der Verteilungen entwickelten Anpassungsmethode. Dieser Teil ist in englischer Sprache verfasst und am 6. August 2009 im Journal "Aerospace Science and Technology" vom Elsevier-Verlag¹ zur Veröffentlichung eingereicht. Eine ausführliche Diskussion der Ergebnisse erfolgt in Kapitel 7 ab Seite 130.

4.1 Abstract

Within the project ICE ("Ideal Cabin Environment", funded by the European Commission under contract no. AST4-CT-2005-516131), the impact of different environmental conditions on comfort was investigated by 1,500 participants in total. The ICE consortium was made up of 15 members from eight countries. Simulated flights of 8 hours each with 1,100 passengers on 29 days were conducted in a large-scale aircraft cabin environment facility near Munich, Germany in winter 2006/2007. During these tests, typical in-flight conditions for noise, vibration, temperature, humidity, and pressure were presented with different level combinations on different days in order to analyze possible effects of these parameters on participants' comfort and well-being. The environmental conditions and passengers' physiological response were monitored by sensors, while psychological response was recorded using PDA-based questionnaires. For our share of evaluation, extensive statistical methods were used, applied on a flight-by-flight analysis.

As one outcome, the possibility to describe a large class of responses by a function based on the Poisson distribution is shown. An approximation of observed data was performed using the d_{max} test criteria of the Kolmogorov-Smirnov goodness-of-fit test. In addition, R as the variance-to-mean ratio (VMR) was determined and compared to the quality of the predictions with a significant equality to the observed

¹http://ees.elsevier.com/aescte, ISSN: 1270-9638, Imprint: Elsevier

data. For subjective ratings on a unipolar scale and a calculated VMR of $R \ge 0.72$ in raw distribution data, a success rate of nearly 95% in representing the observed data distribution with one single parameter was identified. This approach allows for a representation of an average participant's impression (answer distributions of questionnaire data) of a defined environmental condition and enables for a prediction of not measured, interim situations.

4.2 Introduction

An increasing amount of passengers spend more and more time in aircraft cabins [1] and are exposed to environmental conditions created artificially by aircraft's onboard system. Aircraft manufacturers therefore put emphasis on the development of an aircraft cabin interior that offers as much comfort as possible, keeping expenditure in mind. Information on passengers' perception and acceptance of certain environmental conditions as a competitive edge in airlines' choice are of great value. Different methods to measure customer satisfaction are available like market share or turnover of a company as objective and analysis of passengers' complaints or questionnaire surveys as subjective tools [19]. One early investigation of in-flight comfort at short haul travel, using questionnaires with 758 respondents, was performed and published between 1972 and 1978 by Richards and Jacobson [20, 21, 22, 23]. As customer requirements may change with time, several surveys (some of them performed by airline companies and thus unpublished, others performed by research projects, e.g. [7]) followed.

Objective of the ICE project² (scheduled to run for three years from October 2005 to September 2008 and prolonged until March 2009) was to identify an ideal combination of different ambient factors towards agreeableness and comfort for today's flying public. To evaluate combined effects of temperature, relative humidity, cabin pressure, and noise and to provide an European pre-Standard for an optimized aircraft cabin environment [5], a large number of 'passengers' were exposed to different conditions during simulated flights, including not only typical in-flight conditions but also values expected to offer different levels of comfort. Evaluated psychological and physiological tools were used by the consortium to record passengers' perception, effects on health and on well-being as well as passengers' preferences.

The major achievement of the ICE-investigation was to develop a "Predictive

²Official project website: http://www.ice-project.eu

Design Model" (PDM), allowing to predict a passenger's perception of certain environmental conditions in an aircraft cabin and the passenger's satisfaction with those conditions. As a first step, gathered data were analyzed by the authors in order to describe the observed answers (distributions) by an appropriate mathematical approximation. The innovative approach presented in this study should yield a simplified mathematical description of data and thus enable a significant data reduction. Furthermore, the parametrization allows for methods of interpolation of comfort ratings being used in the final ICE-model.

4.3 Data collection

The ICE consortium performed a total of 36 simulated flight tests at the Aircraft Cabin Environment (ACE) rig (7 flights) located at the Building Research Establishment Ltd. (BRE) in Watford, UK [9] and at the Flight Test Facility (FTF) at Fraunhofer Institute for Building Physics (IBP) in Germany [24] (29 flights). Even though in the ICE-project all flights were used in data analysis, the present study aims exclusively at data gathered during FTF simulator-flights. The FTF consists of a fuselage of an Airbus A310 front section with the same interior as used during its previous service. From November 2006 to January 2007, 29 tests took place, each consisting of a 30-minutes pre-base phase, 7 hours of simulated flight and a 30-minutes post-base phase, so the overall time onboard was 8 hours. During "flight", environmental parameters as cabin pressure, relative humidity, temperature, and noise were varied on a day-by-day basis according to predefined test plan. Pre- and post-base phases were held constant for all flights at ambient conditions. For development of the PDM, physiological and psychological data were gathered of approx. 1.150 persons (40 persons per flight), aged from 18 to 93, showing an equal distribution across the age groups 18 to 34, 35 to 50, and 50+ years. Further details on the test design, the parameter variation, and the experimental conditions can be found in |11| and |25|.

During each simulation run in the FTF (divided into 5 segments), a PDA-based questionnaire with 117 CMSB-³ and 93 PC questions⁴ (1x) was handed out to the participants, 3 times during "flight" (segments 2, 3, 4), once during the pre- (segment 1)

 $^{^{3}}$ passenger cabin environment survey

⁴passenger personal characteristics

and once during the post-baseline phase (segment 5). Answers were given on four-, five- or seven-category scales, either unipolar or bipolar. The topics adressed were subjects' state of comfort, mood, symptoms, and behavior (asked at all 5 inquiries) as well as subjects' personal characteristics, state of health, general well-being, and sensitivity to certain environmental situations (asked only once, as these conditions are not likely to change during one "flight"). Questionnaires were developed, supervised, and preprocessed for later analyses by the Unit for Psychology of Ergonomics, Environment, Health and Performance at the Institute for Environmental Hygiene, Medical University of Vienna (MUV), Austria [16].

4.4 Data analysis

During tests at FTF, one flight was aborted due to external circumstances, so 28 flights were available for analysis. After a thorough principal component analysis (PCA) carried out together with MUV, 72 questionnaire items were identified to be part of the future PDM. Numerous items were combined to representative indices.

As a subgroup, a set of 29 questionnaire items and indices was used in this study for approximation approach. 29 items were assigned to five sections, as listed in Table 4.1. A list of all items can be found in the Appendix Table 4.3.

For this listed set, answers were given on a unipolar scale with a proposed prominent cumulation at one end of the scale: For each question, the scale offers a descending sequence of responses from 'no effect', 'no disturbance', or 'optimal condition' to 'severe effect', 'a great deal of disturbance', or 'adverse condition'.

At first, distributions of answers from the 40 participants (per flight) for these questionnaire items and indices were calculated. An example is given in Figure 4.1, showing observed distribution of index $Mean-sym-pain^5$ for flight 5, 1st segment.

⁵*Mean-sym-pain* is a combined index representing the perception of muscle pain in different parts of the body (back, neck, bottom, and legs, based on the results of PCA).

Tabelle 4.1: Overview of the 5 analyzed sections [with abbreviations]. For each section (column 1), quantity of items dealing with this topic and scale, used in the questionnaire, are listed in column 2 and 3. In column 4, an example question is given together with unipolar-scaled response options.

section [abbr.]	consists of	used scale	example quest. / response option		
usage of	1 item	4-cat.	"Have you used a blanket?"		
blankets $[B]$			$1 = \text{not at all } \dots 4 = \text{all the time}$		
environmental	5 items /	4-cat.	"Currently, the air movement is"		
perception $[E]$	indices		$1 = \text{strong} \dots 4 = \text{imperceptible}$		
comfort	8 items $/$	4-cat.(6x)	"For me, the air is currently"		
[C]	indices	7-cat. $(2x)$	$1 = \text{too stale} \dots 4 = \text{comfortable}$		
personal mental	4 items /	7-cat.	"Currently, I feel sleepy"		
state $[st]$	indices		$1 = \text{not at all } \dots 7 = \text{very much}$		
perceived	11 items /	4-cat.	"Currently, I have pain in the back"		
symptoms $[sym]$	indices		$1 = \text{not at all } \dots 4 = \text{very severely}$		



Abbildung 4.1: Example of an observed answer distribution (relative frequency for Flight 5, 1st segment) for combined pain-index **Mean-sym-pain**, asking for *pain in the back / ... neck / ... bottom / ... legs* with response options ranging from "1" (not at all) to "4" (very severely).

Comparing the (right-skewed) shape of observed distributions with well known discrete probability distributions, especially multinomial distribution (as a generalization of the binomial distribution) or the Poisson distribution (described e.g. in [26]), similarities are obvious. Particularly the Poisson distribution has the advantage of describing complete distribution by one single parameter (' λ '). In the following, characteristics of Poisson distribution are presented in detail, as the approximation of the distribution of observed raw data with this model distribution seems most promising.

4.4.1 The Poisson distribution

Let X to be a random variable which can take a number of discrete integer values x with $0 \le x < \infty$, counting the amount of observations during a time-interval of given length. With λ being the expected number of occurrences in this interval, the probability P of observing exactly k occurrences is given by Equation 4.1:

$$P_{\lambda}(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$
(4.1)

where

- $P_{\lambda}(k)$ is probability of k occurrences
- e as base of the natural logarithm
- k is number of occurrences of the event, beginning with k = 0 ($k \in \mathbb{N}_0$)
- λ is a positive real number, representing mean number of occurrences $\langle k \rangle$ and variance Var(X)

4.4.2 Required raw data for further analysis

The discrete Poisson distribution is often used in the area of actuarial science, trying to estimate probability of the number of (mostly rare) events occurring in a fixed period of time (or in other specified intervals like distance, area, or volume) [26, 27].

Although the underlying process of answering a questionnaire is not proven to be poissonian, the shape of observed distributions - as shown in Figure 4.1 - suggests to parameterize the observed distribution by a Poisson distribution with a corresponding λ .

Therefore, as a requirement for a Poisson-like distribution, questionnaire indices with answers given on a unipolar scale with an expected prominent cumulation at one end of the scale are selected for an approximation, as proposed in Table 4.1. In addition, the scale probably has to be flipped over to achieve a positive (right-) skewed distribution in general meaning.

Limitations

Currently, answers with a 'neutral' rating in the middle of the scale (bipolar scale), are not applicable for an approximation by Poisson distribution directly. Preliminary tests with a modified scale were performed with promising results, but these are not part of the present study.

4.4.3 Fitting the observed distribution with a Poisson distribution

To meet the requirements defined in (4.1), the questionnaire data have to be shifted in scale to start at k = 0 if necessary, e.g. an original 4-category scale from 1 to 4 has to be shifted by -1 to obtain a scale from 0 to 3.

Fitting by method of least squares

A common method to approximate observed data by a model function is by adjusting the determinant parameters of the model function in a way to fit observed data best. The least square method (LS) assumes best fit, when the sum of the squared residuals⁶ is at a minimum [28].

The parameter λ as the only parameter of the model function (4.1) was calculated by minimizing the sum of squared residuals of each answer distribution, defined as λ_{LS} in (4.2)

$$\min_{\lambda_{LS}} \left[\sum_{k=1}^{n} \left((P_k(obs) - P_k(\lambda_{LS}))^2 \right) \right]$$
(4.2)

with $P_k(obs)$ as observed and $P_k(\lambda_{LS})$ as predicted probability of answer k, given by the model (Poisson distribution).

⁶A residual is defined as difference between values of observed distribution and values predicted by the model.

4.4.4 Validation of approximation by applying a goodness-of-fit test

The quality of an approximation can be validated by applying statistical tests on the comparison of observed and hypothetical distribution. The *Kolmogorov-Smirnov* D test [29] employs the greatest discrepancy between observed and expected cumulative frequencies d_{max} , as given by (4.3).

$$d_{max} = \sup_{k} \left| \sum_{k_i \le k} P_k(obs) - \sum_{k_i \le k} P_k(\lambda_{LS}) \right|.$$
(4.3)

Test statistic $Z_n = d_{max}\sqrt{n}$, with *n* being the sample size, is then compared with critical value d_{α} (a table of critical values for d_{α} can be found in [30]). For a sample size of n > 35, the critical value can be approximated by $d_{\alpha} \approx 1.36/\sqrt{n}$ at a confidence level of $\alpha = 0.05$ [31, 32].

Test hypothesis

Null hypothesis is given by $\mathcal{H}_0: F = F_0$, with F as observed distribution to be equal to F_0 , the predicted distribution, α and β error are shown in Table 4.2.

Tabelle 4.2: Table of Confusion: α specifies the error of rejecting \mathcal{H}_0 (equality of both distributions) even though they are the same, β gives error of falsely accepting the equality of observed and predicted distribution.

		true condition		
		$F = F_0$	$F \neq F_0$	
test	\mathcal{H}_0 rejected	α		
result	\mathcal{H}_0 accepted		$\beta \leq 1-\alpha$	

Probability p of rejecting \mathcal{H}_0 falsely (α , type I error) is given for the Kolmogorov-Smirnov test by [33, 34]:

$$p_{KS} = 1 - \frac{\sqrt{2\pi}}{Z_n} \sum_{m=1}^{\infty} e^{-(2m-1)^2 \pi^2 / (8Z_n^2)}.$$
(4.4)

So e.g. for $p_{KS} = 0.95$, probability of rejecting \mathcal{H}_0 falsely is $\alpha = 95\%$, indicating that on the other hand, probability β of falsely accepting the equality of observed and predicted distribution is less or equal than 5%.

4.4.5 Alternative method for approximation: minimizing d_{max}

In order to meet the criteria for the Kolmogorov-Smirnov test (given in (4.3)) best, another strategy for approximation may arise by minimizing d_{max} as a boundary condition for deriving λ (= $\lambda_{d_{max}}$):

$$\min_{\lambda_{d_{max}}} \left[\sup_{k} |P_k(obs) - P_k(\lambda_{d_{max}})| \right]$$
(4.5)

Together with λ_{LS} (4.2), two different methods of approximation are available now. Both results of the approximations (Poisson distributions defined by λ_{LS} and by $\lambda_{d_{max}}$) are validated against the observed distribution by the Kolmogorov-Smirnov goodness-of-fit test.

4.5 Results of Kolmogorov-Smirnov goodness-of-fit test (KS-test)

Based on (4.4), the probability p_{KS} of denying \mathcal{H}_0 falsely is calculated twice for each of the 29 items (as defined in Table 4.1), once for the approximation based on λ_{LS} and once for $\lambda_{d_{max}}$. In total, 2 x 4060 tests are performed (28 simulated flights in the FTF with 5 segments each).

4.5.1 Example

The results for the combined seat-comfort index C- $Cond^7$, recorded at flight 5, 3^{rd} segment, are shown in Figure 4.2. In this plot, the distribution is truncated at the upper end $(1 \le k \le 4, \text{ with } k \text{ as the (shifted) response option in the questionnaire).}$

⁷"C-cond" covers four questions concerning general seat comfort, the quality of seat width, of legroom, and the judgement on the proximity.



Abbildung 4.2: Example of an observed answer distribution (answers shifted from 0 - 3 to 1 - 4, as found in the original questionnaire) for the combined seat-comfort index 'C-Cond', taken from Flight 5, segment 3, in comparison to the results of the approximation, performed with two different methods, least squares (LS, left) and minimizing d_{max} (on the right). In addition, the derived λ values and the probability p_{KS} of rejecting \mathcal{H}_0 falsely are shown: $\lambda_{LS} = 1.65, p_{KS} = 0.93$ and $\lambda_{d_{max}} = 1.48$ with $p_{KS} = 1.00^{**}$, confirming not to reject equality of observed and predicted distribution (by $\lambda_{d_{max}}$).

Application of different approximation methods will obviously result in slightly varying λ -values. The least square fit leads to a λ_{LS} of 1.65 and $\alpha = 93\%$. With $\beta \leq 1 - \alpha = 7\%$, one can't be sure to reject alternative hypothesis of observed and assumed distributions being unequal (\mathcal{H}_1) on a confidence level of 5%. On the other hand, for $\lambda_{d_{max}} = 1.48$, obtained by minimizing d_{max} , $\beta < 1\%$ and one can be sure not to falsely accept equality of distributions.

4.5.2 Analysis of sections

Frequency of probabilities p_{KS} with $p_{KS} \ge 95\%$ (indicating that equality between observed and predicted distribution cannot be rejected on a level of $\le 5\%$) is given (in %) now for each questionnaire section (Figure 4.3). Again, results are divided into two groups, λ_{LS} and $\lambda_{d_{max}}$, based on the two different approximation methods.



Index-domain

Abbildung 4.3: Frequency [in %] of probabilities p_{KS} with $p_{KS} \ge 95\%$ for all investigated questionnaire sections, confirming an error probability of $\beta \le 5\%$, gained from two different methods of approximation: method of least squares (*LS*, all left bars) and by minimizing d_{max} (d_{max} , right bars respectively). Predicted distributions are based on the equation for the discrete Poisson distribution.

For the group of all investigated questionnaire items and indices (Figure 4.3) and for five subsections (defined in Table 4.1), approximation by the method of least squares (*LS*, all left bars in red) leads to a lower amount of significant p_{KS} -values and the approximation by minimizing d_{max} (all right bars) to a higher amount. Furthermore, the section of perceived symptoms shows significant results in nearly all cases, whereas the questionnaire domain dealing with environmental perception (*E-questions*) shows significant equality for only about 40% of the tests.

Though analyzed questionnaire items and indices are well selected based on requirements defined in chapter 4.4.2, results shown in Figure 4.3 suggest to introduce another criterion for a (pre-) selection of questions to be approximated. By its definition (see chapter 4.4.1), the one-parameter Poisson distribution allows for a simple pre-test by comparing mean and variance of an observed distribution, as described in the following.

4.5.3 Variance-to-mean ratio (VMR)

Calculating ratio R of variance to the mean for every single observed set of answers, three possibilities can occur:

- R < 1: distribution is considered *under-dispersed*,
- R = 1: distribution is *equally dispersed*, as defined for Poisson distribution,
- R > 1: distribution is considered *over-dispersed*.

Figure 4.4 shows the calculated VMRs of all observed questionnaire items and indices (N = 4060).



Abbildung 4.4: Histogram of calculated variance-to-mean ratios (VMR) for all questionnaire items and indices of all 28 flights, performed in the FTF (N = 4060). With R < 1, distributions are considered to be underdispersed ($\approx 81\%$ of all observed distributions).

About 81% of distributions are considered to be underdispersed with R < 1. Combining information on probabilities of $p_{KS} \ge 95\%$ (Figure 4.3) and spreading of VMRs (Figure 4.4), a predication can be done on the dependency between these two parameters (Figure 4.5). The relative number of significant results [in %] was determined in the range of $0.05 \le R < 2.95$ for each interval integrated over an interval length of $\Delta R = 0.1$.



Abbildung 4.5: Dependency between calculated VMR of observed data (x-axis, for each interval length of $\Delta R = 0.1$) and frequency of those KS-tests, which exhibit a significant result in the corresponding interval (yaxis, [in %]). For visualization, discrete results (marked with boxes) are connected with a dotted line. For a VMR with $0.7 \le R \le 1.2$ the percentage of significant p_{KS} -values is > 90%.

For $0 \leq R < 0.4$, the validation by the KS-test shows no significant result for the approximations. Between R = 0.4 and R = 0.7, the percentage of significant results is increasing noticeably. For $R \geq 0.72$, more than 95% of all p_{KS} -values are significant. This number of significant values is then slightly decreasing to 75% for VMRs with $1 < R \leq 2.1$. For R > 2.1, nearly all approximations are significant with $p_{KS} \geq 95\%$ (with only a low number of cases, as shown in Figure 4.4). In the following, for each section (defined in Table 4.1), a histogram of p_{KS} -values (left side) and the relation between VMRs and derived p_{KS} -values (on the right side as scatter plot) are given. This representation allows for section-by-section analysis of this dependency.

4.5.4 Results per section

Result of section "behavior"

Section behavior consists of only one question, meeting requirements for an approximation by a Poisson distribution as defined in chapter 4.4.2. 140 cases are tested (28 flights, with 5 segments). Figure 4.6 (left) shows the derived p_{KS} -values by the KS-test as a histogram. On the right side, relation between variance-to-mean ratio and corresponding p_{KS} -value is shown. The horizontal line indicates the 95%-level, the vertical line marks the R = 1 position, defining an equally dispersed distribution.



Abbildung 4.6: Result of section "behavior": histogram of p_{KS} -values (left) and relation between VMR and p_{KS} -value (right) for one item, the usage of blankets with the response option 1 = not at all ... 4 = all the time. In addition, a horizontal line indicates the 95%-level and a vertical line at R = 1 marks the position of an equally dispersed distribution.

For about 84% of all p_{KS} -values, $p_{KS} \ge 95\%$, indicating an error probability of $\beta \le 5\%$. Observed distributions are mainly considered to by over-dispersed, as R > 1 for approx. 75% of all cases.

Results of section "environmental perception"

For this section, 5 items are analyzed (700 cases in total, based on 28 flights, 5 segments and 5 items). The left side of Figure 4.7 shows derived p_{KS} -values by the KS-test as a histogram. On the right side, relation between VMR and corresponding p_{KS} -value is shown. Again, a horizontal line marks the 95%-level and a vertical line at R = 1 marks the position of an equally dispersed distribution.



Abbildung 4.7: Results of section "environmental perception": histogram of p_{KS} -values (left) and relation between VMR and p_{KS} -value (right).

As shown in Figure 4.7 left, only 40% of derived p_{KS} -values show an error probability of $\beta \leq 5\%$. Observed distributions for this section are mainly considered to by underdispersed, as R < 1 for approx. 97% of all cases.

Results of section "comfort"

For "comfort" section, 8 items are analyzed (1120 cases), two of them using a sevencategory scale.



Abbildung 4.8: Results of "comfort" section: histogram of p_{KS} -values (left) and relation between VMR and p_{KS} -value (right).

As shown in Figure 4.8, an amount of about 75% of the derived p_{KS} -values show an error probability of $\beta \leq 5\%$. Here, observed distributions are mainly considered to by underdispersed, as R < 1 for approx. 94% of all cases.

Results of section "personal mental state"

Section "personal mental state" consists of four different items on a seven-category scale in deviation to other (four-category scaled) items. 560 cases are tested in total (see results in Figure 4.9).



Abbildung 4.9: Results of section "personal mental status": histogram of p_{KS} -values (left) and relation between VMR and p_{KS} -value (right).

For section "personal mental state", an amount of 63% of derived p_{KS} -values show an error probability of $\beta \leq 5\%$. These distributions are mainly considered to be under-

dispersed, but also partly over-dispersed ($\approx 27\%$). About 6% are equally dispersed with R = 1.

Results of section "perceived symptoms"

For section of "perceived symptoms", 11 items are analyzed in 1540 cases (Figure 4.10).



Abbildung 4.10: Results of section "perceived symptoms": histogram of p_{KS} -values (left) and relation between VMR and p_{KS} -value (right).

Here, an amount of about 99% of derived p_{KS} -values show an error probability of $\beta \leq 5\%$. The VMR-values are mainly distributed between R = 0.5 and R = 1.8 with a maximum at R = 0.9, indicating a nearly equally dispersed distribution.

4.5.5 Summary

For the question of section 'b' (usage of blankets) and most of the questions dealing with 'perceived symptoms' ('sym'), the proposed approximation by the equation of the Poisson distribution is in good accordance with raw data distribution. For the section of "environmental perception" with R < 1 for approx. 97% of all investigated cases, only 40% of the fitting are in line with observed distributions.

Based on preliminary considerations presented in chapter 4.5.3, the number of p_{KS} with $p_{KS} \ge 95\%$ is observed to be greatest at $R \ge 0.72$. Comparable to Figure 4.3, the frequency of probabilities p_{KS} with $p_{KS} \ge 95\%$ is given (in %) for each questionnaire section in Figure 4.11. For the approximation method d_{max} , results

are given for distributions with $R \ge 0$ and for those distributions with $R \ge 0.72$ in the observed data.



Index-domain

Abbildung 4.11: Comparable to Figure 4.3, the frequency [in %] of probabilities p_{KS} with $p_{KS} \ge 95\%$ for all investigated questionnaire sections, confirming an error probability of $\beta \le 5\%$, are shown (approximated by the d_{max} -method). Results are given for all distributions (with $R \ge 0$, all left bars) as well as for those distributions with $R \ge 0.72$ in the observed data (all right bars).

As shown in Figure 4.11, results for each section (relative frequency of p_{KS} with $p_{KS} \ge 95\%$) can be improved for distributions with a VMR of $R \ge 0.72$, except for the section 'behavior', as all distributions in this sections already show a VMR of R > 0.72.

4.6 Application: Predicted Percentage of Dissatisfied - PPD

4.6.1 Definition

In 2004, the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) defined the "ANSI/ASHRAE Standard 55-2004", dealing with the "Thermal Environmental Conditions for Human Occupancy" [35]. A thermal sensation scale (which allows for quantifying people's thermal sensation) is defined as follows:

+3 = hot, +2 = warm, +1 = slightly warm, 0 = neutral,-1 = slightly cool, -2 = cool, -3 = cold

The predicted percentage of dissatisfied (PPD) is then based on the assumption that people voting +3, +2, -2, or -3 on the thermal sensation scale are dissatisfied.

In this study, "predicted percentage of dissatisfied" (or for perceived symptoms: "predicted percentage of people suffering from a perceived symptom") is derived for both data sets (distributions), observed PPD (PPD_{obs} , based on questionnaire data) and predicted PPD, based on d_{max} approximation method (chapter 4.4.5), $PPD_{d_{max}}$, by (4.6) and (4.7) for the four-category scales.

$$PPD_{obs} = 1 - \sum_{k=1}^{2} P_k(measured) \tag{4.6}$$

or

$$PPD_{d_{max}} = 1 - \sum_{k=1}^{2} P_k(\lambda_{d_{max}})$$
(4.7)

Thus PPD is based on those participants, voting 3 (= distinctly) or 4 (= very severe)⁸ on the unipolar four-category scale or 4, 5, 6, 7 on an unipolar seven-category scale with k = 1, 2, 3.

4.6.2 Quality of PPD prediction

Comparison of predicted $PPD_{d_{max}}$ and observed PPD_{obs} is shown in Figure 4.12 for two cases as scatter plots. Left, all analyzed 4060 cases (for 28 flights, 5 segments,

⁸e.g. for perceived symptoms

29 indices) are plotted. On the right side, only those PPDs with a VMR of $R \ge 0.72$ in the observed data are compared, based on the results of chapter 4.5.5.



Abbildung 4.12: Predicted and observed PPD as scatter-plot. On the left: for all analyzed questionnaire items and indices $(R \ge 0)$. On the right: truncated data set (only those items and indices with a VMR of $R \ge 0.72$). For both, the ideal state with $PPD_{obs} = PPD_{d_{max}}$ is shown as a thin line. In addition, a linear regression is performed (thick line).

Though the ideal prediction of a PPD with $PPD_{obs} = PPD_{d_{max}}$ or y = x is not fully met, forecast for the reduced data set (with $VMR \ge 0.72$) with $y \approx 1.09x$ as result of a linear regression is significant (r = 0.98, p < 0.05). For the complete data set, the linear regression is given by y = 1.05x - 0.02 with r = 0.95.

4.7 Conclusion and outlook

Within the European project ICE, 36 simulated flights were performed and questionnaire data of about 1500 'passengers' were collected to develop a predictive design model, allowing to predict passengers' perception of certain environmental conditions in an aircraft cabin and their satisfaction. Based on a thorough principal component analysis, a subset of 72 questionnaire items and indices was derived. The items and indices with an expected prominent focus on one end of the unipolar scale (all in all 29 items and indices) were parameterized by fitting the observed data with the equation of the Poisson distribution (the possibility to approximate answer distributions on a bipolar scale is not yet fully investigated and therefore not part of this study). For approximation, two different methods were applied on observed data gathered at simulated flight tests at the Fraunhofer Flight Test Facility FTF: method of least squares and, based on the goodness-of-fit criterion of the KS-test, method of minimizing the maximum difference d_{max} between cumulated distribution function of raw data and predicted data. The quality of approximation was then tested by the Kolmogorov-Smirnov goodness-of-fit test. Based on the number of observed probabilities with a β -error of $\beta \leq 5\%$ ($p_{KS} \geq 95\%$) suggesting to reject the alternative hypothesis \mathcal{H}_1 of inequality of both distributions, the method of minimizing d_{max} is more successful.

In addition, with calculation of R as the variance-to-mean ratio (VMR), a criterion for a successful application of proposed approximation was identified: For observed answer distributions with a VMR of $R \ge 0.72$, about 95% of all tested approximations in this study were significant on the 5% error level.

Finally, with the described possibility to parameterize each of the observed distributions of questionnaire data with one parameter (λ) , gained value can be assigned to a corresponding environmental condition offered during the simulated flight. As a result of an applied (linear) regression, a prediction can be derived for interim values, not directly tested during the original test campaign.

4.8 Acknowledgement

The authors are very grateful to Prof. Dr. Volker Mellert and Dr. Reinhard Weber for their support and helpful discussions. We would also like to thank Dr. Roland Kruse and Daniela Poppinga for assistance with the statistical analysis.

The European commission financially supported this work under contract number AST4-CT-2005- 516131. The authors wish to thank the consortium consisting of the following partner organizations:

- BRE, United Kingdom, Project Coordinator
- Airbus Deutschland GmbH, Germany
- Antanas Gustaitis Aviation Institute of Vilnius Gediminas Technical University, Lithuania
- Avitronics Research, Greece
- Civil Aviation Authority, Aviation Health Unit, United Kingdom
- Deutsches Zentrum für Luft- und Raumfahrt e. V., Germany
- EADS Innovation Works, Germany
- Fraunhofer Institut für Bauphysik, Germany
- Medical University of Vienna, Austria
- Stichting National Lucht en Ruimtevaartlaboratorium, The Netherlands
- Royal Free and University College Medical School, United Kingdom
- Streit Technische Gebäude Ausrüstung, Germany
- Carl-von-Ossietzky Universität Oldenburg, Germany
- Universita delgi Studi di Padova, Italy
- Czech Technical University in Prague, Czech Republic

4.9 Appendix

Tabelle 4.3: Overview of analysed items and indices (coloumn 1, number of combined questions in brackets) and corresponding topic (coloumn 2). Used scales are four-category, except those which are marked (*), offering a seven-category scale. The response options are shown in last coloum.

		response option		
$\mathbf{index}\ /\ \mathbf{item}$	topic	\mathbf{best}	worst	
B_blanket (1)	usage of blanket	never	always	
$E_{draught}$ (1)	perception of air movement	imperceptible	strong	
$E_{stale}(1)$	perception of air staleness	not at all stale	very stale	
$E_{odour}(1)$	perception of odours	imperceptible	strong	
$E_{passenger}(1)$	noises from passengers	none	many	
$E_{vibration}(1)$	perception of vibration	imperceptible	strong	
C_airstale (1)	rating of air staleness	comfortable	too stale	
C_{air} (3)	overall air comfort index	$\operatorname{comfortable}$	annoying	
C_odour (1)	rating of odours	$\operatorname{comfortable}$	annoying	
C_noise (2)	overall noise comfort index	$\operatorname{comfortable}$	annoying	
C_vibration (1)	rating of vibrations	$\operatorname{comfortable}$	uncomfortable	
C_cond (4)	overall seat comfort index	$\operatorname{comfortable}$	uncomfortable	
$C_condition$ (4)	overall comfort index $(*)$	$\operatorname{comfortable}$	uncomfortable	
$C_{wellbeing}$	overall well-being index $(*)$	very good	very bad	
$St_fatigue (3)$	feeling sleepy/tired (*)	not at all	very serverly	
$St_g_{mod}(3)$	feeling happy $(*)$	very much	not at all	
St_concentration (3)	ability to concentrate $(*)$	very much	not at all	
St_apathetic (3)	feeling bored $(*)$	not at all	very serverly	
$sym_pain(5)$	general pain	not at all	very serverly	
$sym_dryness$ (4)	dryness symptoms	not at all	very serverly	
$sym_headache (3)$	suffer from headache	not at all	very serverly	
sym_flu (4)	have a cold	not at all	very serverly	
sym_stress (3)	being stressed out	not at all	very serverly	
$sym_{freeze}(3)$	have cold feet/hands	not at all	very serverly	
$sym_stomach(3)$	problems with the stomach	not at all	very serverly	
sym_eyes (2)	eyes related symptoms	not at all	very serverly	
sym_tired (1)	being tired	not at all	very serverly	
$sym_pressure(2)$	effects from low pressure	not at all	very serverly	
Sym_average (36)	overall symptom index	$\mathrm{none}/\mathrm{few}$	a lot	

5 ICE IV – Investigation of temporary threshold shifts in aircraft cabins

Das 5. Kapitel dokumentiert die im ICE-Projekt eingegliederte Untersuchung zur möglichen temporären Hörschwellenverschiebung (Temporary Threshold Shift – TTS) nach einer mehrstündigen Schallexposition in einer simulierten Flugzeugkabinenumgebung. Dieses Kapitel ist ebenfalls in englischer Sprache verfasst und wird zur Veröffentlichung eingereicht. Eine weiterführende Diskussion der Ergebnisse erfolgt in Kapitel 7 ab Seite 138.

5.1 Introduction and theoretical background

The mobility of humans has been increasing considerably during the past decades. Especially in air travel during the last 10 years the number of passengers has mounted by about 5% per year and this growth is also expected in future [1]. Hence an increasing amount of passengers will spend longer time periods in aircrafts being exposed to cabin specific environmental parameters. It is the goal of the EU project "Ideal Cabin Environment" (ICE, www.ice-project.eu) to find an optimal combination of the environmental parameters as temperature, relative humidity, noise, and pressure that guarantee the passengers comfort on long haul flights. As a scientifically desirable alteration of values of environmental factors is not possible during real flights, comfort investigations have been performed in the frame of the project in a ground-based Airbus A310 cabin simulator (flight test facility – FTF¹). Besides temperature, relative humidity and pressure, noise level was one variable to be changed in the parameter configuration to be tested for comfort. During different 7 hours of simulated flights cabin background noise, imitating normal cruise noise, was set to 64 dB(A), 69 dB(A) and 74 dB(A)². Partly, these levels are already realized in modern aircrafts, partly they are aspired in future cabins. Concerning passenger's as well as crew's comfort and health it is of specific interest to know

¹See http://www.bauphysik.de/ftf/index_e.html, July 2009, for details

 $^{^{2}}$ see also [25, 11, 5] for further details of the ICE project

whether those noise levels may lead to physiological impairment of hearing system in terms of temporary threshold shifts.

High noise exposure in everyday life [36] may cause an impairment of human auditory system, apart from biological aging process. The impairment can affect human hearing threshold twofold that can be classified either as a permanent shift of hearing threshold (PTS) or as a time limited (temporary) threshold shift (TTS). TTS is defined as a temporary degradation of the individual hearing threshold as a characteristic response of auditory system [37] to high-level noise exposure. Biologically the temporary desensitisation of hearing organ is understood as a consequence of a metabolic fatigue [38, 39].

Experiments investigating TTS usually comprise three sections [37]. First section: before an exposure ("pre-exposure"), the normal unaffected individual hearing threshold $(L_{T,pre})$ is measured in a sound proof cabin. Second section: the participant is exposed to fatiguing noise stimuli for a defined period of time ("exposure"). Third section: during post-exposure section, individual hearing threshold $L_{T,post}$ is determined at the end of specific time intervals after end of the exposure period. TTS decreases as the time lapse after exposure increases, which makes it necessary always to indicate at which instant after exposure the TTS is measured. E.g., TTS after 2 min is denoted as TTS_2 and TTS_{30} is measured 30 min after exposure time.

The magnitude of a TTS is defined as the difference between hearing threshold level $L_{T,post}$ after and threshold $L_{T,pre}$ before noise exposure by

$$TTS = L_{T,post} - L_{T,pre} \tag{5.1}$$

Concerning the parameters influencing the threshold shift, Moore [40] identified the following factors: frequency distribution of fatiguing stimulus, its duration and intensity as well as the time period between end of exposure and point of time of threshold determination.

Regarding *frequency* of the fatiguing stimulus, following dependencies were observed. For tonal stimuli and narrow band signals above 250 Hz, many authors found a main threshold shift about 1/2 octave above centre frequency of the stimulus [41, 42, 43, 44, 45, 46, 47, 48, 49], among others). Broad band signals mostly induce a maximum threshold shift in most sensitive frequency area of the auditory system (3 to 5 kHz) [50] or 3 to 6 kHz [47]. For low frequency exposures, knowledge is limited. Mills *et al.* (1983) [51] reported a TTS not necessarily 1/2 octave above stimulus frequency, but in the region of higher sensitivity. In many studies TTS is only determined at that frequency that shows highest magnitude of threshold shift.

The *duration* of exposure is decisive for the growth of TTS during and recovery after exposure. Mills *et al.* (1979) [46] reported different forms of TTS time courses, depending on exposure level (see Figure 5.1) whereas Melnick [50] alluded to a simple exponential function for growth and decay. The maximum value of a TTS was observed by Mills *et al.* [46] between 16 and 24 hours of exposure. With ongoing exposure, TTS saturates and does no longer increase.



Abbildung 5.1: TTS_4 -values after Mills *et al.* (1979) [46], measured after exposures with different levels as a function of exposure duration (fatiguing noise: octave-band noise, centre frequency 4 kHz, measurement frequencies 3.5 kHz resp. 6 kHz).

The *intensity* of fatiguing stimulus is identified as another moderator for growth and the amount of a threshold shift (e.g., Figure 5.1 [46]; also [47, 52, 53] and others). Regarding magnitude of TTS, an 8 hour exposure of 75 dB, e.g., is reported to produce a temporary threshold shift (TTS_4) of about 2 to 3 dB, whereas an 80 dB stimulus leads to a TTS_4 of about 8 dB (83 dB: TTS_4 of about 15 dB). The hypothetical growth as a prediction of a threshold shift measured 2 minutes after a single and continous exposures to noise (TTS_2) is described by Miller [54] in Figure 5.2. These threshold shifts are predicted for average, normal hearing young adults exposed to a band of noise or pure tone centered near 4 kHz. Individual results may vary. For comparison reason to Figure 5.1, additional marks are inserted at an 8 hour exposure: For an exposure of 75 dB a TTS_2 of about 12 dB is expected (for 80 dB: $TTS_2 \approx 20$ dB). The expected TTS_2 -values are 2 to 4 times higher than TTS_4 -values measured by Mills *et al.* (1979) [46].



Abbildung 5.2: Hypothetical growth as a prediction of a threshold shift measured 2 minutes after a single and continous exposures to noise (TTS_2) [54]. Additional marks (arrows) are inserted at an 8 hour exposure: $TTS_2(75 \text{ dB}) \approx 12 \text{ dB}$ and for 80 dB: $TTS_2 \approx 20 \text{ dB}$.

The *time period* between end of exposure and point of threshold determination (time of regeneration or restitution) also proofs to be relevant for the magnitude of

a TTS. Recently Strasser (2006) [55] measured hearing threshold shifts after 1 or 2 hour exposures of 94 dB(A) / 91 dB(A) at different times and fitted the results with a decreasing exponential function (see Figure 5.3). Different recovery curves of TTS were found according to the stimuli (different kind of noises) used.



Abbildung 5.3: Different regression functions as a forecast for a recovery after a temporary hearing threshold shift (based on recent studies of Strasser (2006) [55]).

Kuronen [52] measured the TTS_{30} of military pilots after a 60 or 90 minutes flight with an exposure level of 87 to 91 dB(A) during flight and reported an average threshold shift of about 1.35 dB. This is below the results, reported by Strasser (2006) [55], who found for white noise a TTS_{30} of 5.2 dB on average.

However, no data on TTS of passengers in aircraft cabins are available up to now. Regarding noise level in aircraft cabins, noise measurements within the HEACEproject [8] yielded average sound pressure levels between 70 dB(A) and 84 dB(A). Bagshaw and Lower [56] found levels between 70 and 77 dB(A) and in helicopters, noise levels between 90 and 118 dB(A) were observed by Küpper *et al.* [57].

Objectives

The aim of this paper is the investigation of potential temporary threshold differences ΔL_T in aircraft cabins during long-haul flights. Due to the possible control of noise environment, these investigations are carried out before and after simulated flights within the ICE-project.

To measure threshold differences ΔL_T of air travel passengers, the audiometric tool to determine the hearing threshold had to met some special requirements. Because noise levels in modern aircraft cabin are far below 85 dB(A) (max. L_{A,eq}-level that is accepted at a workplace [58]) only small temporary threshold differences are expected, if at all. In this context a highly accurate, effective and rapid threshold measurement is needed. Regarding an operation before and after noise exposure in a realistic flight simulation environment, the procedure has to be portable, simple and robust in its handling and should, because of a rapid restitution, achieve a result for several measurement frequencies in short time (maximum of 3 minutes, incl. repeated measurements max. 10 minutes). This allows keeping additional burden on volunteers low.

In the course of the project a threshold measurement system, meeting the requirements mentioned, has been developed, implemented on laptops (chapter 5.2). In order to be able to identify possible temporary hearing threshold shifts (TTS) in an aircraft cabin, two additional acoustic environments had been tested. One low noise environment, an office, where no TTS has to be expected. In this environment, the statistics of normal hearing threshold variation within a day have been tested in order to be able to discriminate statistical fluctuations from bias that may be identified as TTS. These investigations, serving at the same time as suitability test for the developed fast threshold measuring system are described in chapter 5.3.

The second acoustic environment consisted of noise, spectrally shaped as aircraft cabin interior noise, at a level of 80 dB(A) presented during 7 hours. Exposure of this level is expected to induce a temporary threshold shift [46]. Chapter 5.4 gives the details and results of this test. Finally, the description of hearing threshold measurements within the ICE project before and after 18 simulated flights with different noise levels is given in chapter 5.5. A summary is given in chapter 5.6 and chapter 5.7 concludes the results.

5.2 Development and testing of a fast audiometric tool for threshold difference measurements

The audiometric tool presents a test tone via headphone continuously increasing in level with certain rate of change, until the participant presses a control button as a sign she/he perceives a tone. As soon as the tone becomes inaudible, communicated by the participant by pressing again the control button, the direction of level change is reversed and the following level decrease ends. Main task here is to keep the signal level close to the individual hearing threshold with a decreasing rate of change. This method is based on a standard Békésy-tracking method [59] with an adaptively decreasing rate of level-change. A schematic cut of the procedure is given in Figure 5.4.



Abbildung 5.4: Schematically cut of modified Békésy-procedure (based on [59]) for one measurement frequency (i = 1), used for determination of individual hearing thresholds. Signal level is changed step-wise by ΔL adaptively up to fourth reversal point. The arithmetic mean L_{T_i} (see Equation 5.2) of the last four reversal points (L_{rev_5} to L_{rev_8} , with $\Delta L = 1$ dB) is stored as designated threshold value for this test tone frequency.

In comparison to the original Békésy-tracking, this method allows to approach the

listener's sensitive level range around hearing threshold very quickly after very few steps due to large level changes in the beginning followed by smaller steps (with a reduced rate of level-change) near the circumnavigated threshold.

5.2.1 Implementation of the audiometric tool

The system is realised on a commercial notebook with an external sound card (M-Audio Fast Track Pro, stereo up to 24 Bit / 96 kHz) and test tones are presented via closed headphones (Sennheiser HDA-200) (see Figure 5.5 for a block diagram of setup). It is controlled by a Matlab-based software routine, realised by Ping Rong, a student of Oldenburg University. The set-up was calibrated with Bruel&Kjäer Ear Simulator TYPE 4157.



Abbildung 5.5: Set-up of audiometric tool, based on a notebook, an external M-Audio sound card and a Sennheiser HDA-200.

5.2.2 Stimuli and presentation

The hearing thresholds are measured for a predefined set of following centre tone frequencies 0.25, 0.5, 1, 2, 4, and 6.3 kHz, one after another, starting with the lowest frequency. The test stimulus itself is a sine wave with a modulation of frequency of $f_{mod} = 15$ Hz around centre frequency, a so called 'wobbletone'. The modulation width is set to $1/10^{th}$ octave and therefore frequency dependent. Using such a frequency modulated test tone will avoid measuring a fine structure of basilar membrane, which may exists (as described in literature [60, 61]). The selection of $f_{mod} = 15$ Hz is based on internal pre-tests, in which this modulation has been assessed neither as "restless" nor as "soporific".



Abbildung 5.6: Scheme of three pulse levels of pulse train used for determination of individual hearing. The tone level consecutively changes step-wise by ΔL (here increasingly), initially starting with 8 dB and then decreasing to 4, 2 and 1 dB after third reversal point (see also Figure 5.4). The presentation of stimuli are adapted from original Békésy-tracking.

The wobbletone is presented discontinuously in form of pulse trains. Tone pulses with a duration of 350 ms each are separated by pauses (150 ms) and consecutive pulses are either increased or decreased by a certain level difference ΔL . Figure 5.6 shows a schematic drawing of three pulses of pulse train in case of a level increase by ΔL for adjacent tone pulses. Starting level of a threshold measurement of one frequency is always clearly below threshold. Then, level increases stepwise with an initial step size of $\Delta L = 8$ dB (divided into 2 steps of 4 dB within 1 second) until the listener detects the test signal and presses a key. This level at first reversal point is denoted by L_{rev_1} .

A subsequent reduction of level with a step size of $\Delta L = 4 \text{ dB}$ (again divided into 2 steps of 2 dB within 1 second) is repeated until the volunteer indicates, again by a key-click, that he/she does no longer perceive the signal (second reversal point, L_{rev_2}).

The listener sees two buttons in front on laptop screen one after the other. One is labeled: 'Please press this button once as soon as the tone is audible' and second 'Please press this button once as soon as the tone is just no longer audible'. The first button with the original label is shown in the Appendix, Figure 5.26. An intervention by the investigator in charge is not necessary; the determination of individual hearing threshold is carried out completely self-contained.

After a total of 8 reversal points the hearing threshold is determined in that way with a rule driven reduction of ΔL from $\Delta L = 8$ dB over 4 dB to 2 dB and $\Delta L = 1$ dB for the last 4 reversal points, as shown in Figure 5.4.

Betke [62] has shown that arithmetic mean of levels at upper and lower reversal points can be used, equally weighted, to designate the hearing threshold. Consequently here the arithmetic mean of last four reversal points is stored as designated threshold value L_{T_i} for this measurement, where it is used as a label for the repetitive measurement for this tone frequency. It is given by Equation (5.2).

$$L_{T_i} = \frac{1}{4} \sum_{k=5}^{8} L_{rev_k}, \tag{5.2}$$

with i = 1, 2, 3 as number of repeatitions for this frequency and k indicating the number of reversal points (see Figure 5.4).

5.2.3 Measurement protocols

Two different measurement protocols HTM1 and HTM2 have been used throughout the experiments (Figure 5.7). In cases during testing phase and field test with enough time for threshold determination, protocol HTM1, an extended arrangement of test tone frequencies, is chosen different to measurement protocol HTM2 in time critical situations.



Abbildung 5.7: Two different measurement protocols with frequency orders will be used for hearing threshold measurement (HTM): measurement protocol HTM1 is a complete run of measurements, also described in Table 5.1. Here, the threshold of every frequency will be determined three times. In measurement protocol HTM2, thresholds of six frequencies are measured only once.

HTM1, designed as extended arrangement, allows to measure thresholds of test tone frequencies three times. The order of test tones from low to high frequencies is essentially preserved. This arrangement consists of three subsets of test tone frequencies that are schematically represented in Figure 5.8.



Abbildung 5.8: Schematic diagram of used frequency blocks FB1 to FB3 for threshold determination over time: FB1 offers at first a pair of two training runs at 0.25 and 500 Hz, which results will be omitted in the calculation of the mean. Second frequency block FB2 shows a complete set of six test frequencies. In block FB3, the two frequencies 250 and 500 Hz are measured.

First set of test tones (FB1) comprises all frequencies in ascending order. But only thresholds of last four frequencies (1, 2, 4 and 6.3 kHz) are used for later analysis. First two test tones with frequencies of 250 Hz and 500 Hz only serve for training

purposes. They are contained in subset 3 (FB3) and will be repeated at the end of the threshold measurements. Subset 2 (FB2) comprises all frequencies as subset 1 but this time all measured thresholds are used in further exploitation. The whole arrangement of test tone frequencies used during testing phase is shown in upper part of Figure 5.7: subset 1 is followed by twice the subset 2 and is terminated by subset 3. Table 5.1 gives a detailed overview on test tones.

For time critical tests, the arrangement of test tone frequencies of measurement protocol HTM2 is shown in the lower part of Figure 5.7. Here, thresholds of all six test tones are measured once (subset 2 of Figure 5.8) in ascending order of their frequencies.

Tabelle 5.1: Summary of one complete set of measurement (measurement protocol HTM1, consisting of $1 \times FB1$, $2 \times FB2$ and $1 \times FB3$, as shown in Figure 5.7)

0.1).							
Measurement freq. f	250 Hz	$500 \; \mathrm{Hz}$	1 kHz	$2 \mathrm{kHz}$	$4 \mathrm{kHz}$	$6.3 \mathrm{kHz}$	
modulation f_{mod}	15 Hz						
modulation width	1 / 10 octave						
duration of signal	$350 \mathrm{~ms} + 150 \mathrm{~ms}$ pause						
Number of threshold							
measurements per	4, [incl. 1x training] 3						
frequency i							
duration time DT							
[+ training sequence]	$DT = \sum_{i=1}^{3} \sum_{f=250 \text{ Hz}}^{6300 \text{ Hz}} DT_{i_f} \left[+ \sum_{f=250 \text{ Hz}}^{500 \text{ Hz}} DT_{1_f} \right]$						
for a $HTM1$ -run				2		-	

5.3 Hearing threshold fluctuations in a non-TTS causing environment

5.3.1 Question

As threshold shifts in aircraft cabin are expected to be small because of the moderate noise levels, it is of special interest to know the amount of threshold variations during a "normal" day in a rather quiet environment, causing no TTS. So the statistics of hearing threshold variations were studied in an office environment by performing hearing threshold measurements in the morning before and in the evening after office work day.

5.3.2 Exposure

Actual noise exposure was measured by a Brüel&Kjaer noise dosimeter on three different days during work time in usual office workplace. The dosimeter was worn each day by one of the eight participants, resulting in a time weighted average level $L_{eq.8h}$ of 50.5 dB (day one: 53.8 dB, day two: 48.8 dB, and day three: 48.9 dB).

5.3.3 Participants

In this study eight³ (self reported) normal hearing participants took part. Three females and five males aged between 25 and 39 years.

5.3.4 Procedure

On three (not mandatory consecutive) days (d = 1, 2, 3), hearing thresholds of participants were measured in addition to their common daily routine at the beginning of the work day and after $6\frac{1}{2}$ hours again (on average). These two points in time are denoted by the variable *time of day*, abbreviated *tod*, which takes the value $tod = tod_{pre}$ (or tod = morning) for the beginning of work day prior to exposure and $tod = tod_{post}$ (or tod = evening) for the end of work day. The routine is sketched in Figure 5.9.

Hearing thresholds were always determined in a separate anechoic room, using the extensive threshold testing program (HTM1). Besides tests on stability of programmed routine the durations of measurement sequences (described in Figure 5.7 and Table 5.1) were monitored.

5.3.5 Results

In the following, different aspects of threshold fluctuations are analysed. After presenting the observed individual hearing thresholds L_T of each participant p in

³In literature, threshold investigations were performed with a varying amount of participants. In most studies, between 4 and 12 participants took part [41, 42, 49, 63, 64, 65, 66, 67, 68]. Pilot studies however used mainly one participant (e.g., [43, 45, 69]).



Abbildung 5.9: Daily routine when measuring differences between 'morning' and 'evening' hearing thresholds. Extended threshold determinations have been performed both times (test program 'HTM1' is outlined in previous chapter). Noise exposure during office work day was 50 dB(A) on average, captured by means of a dosimeter that was worn on three different days by one of the eight participants.

chapter 5.3.5, the calculation method for determination of threshold level difference ΔL_T is given in chapter 5.3.5, followed by an analysis of frequency dependent threshold differences ($\Delta L_T(f, p)$, chapter 5.3.5). Then, a day-dependent threshold difference analysis ($\Delta L_T(p, d)$) is given in chapter 5.3.5, followed by an investigation of standard deviation σ_T in chapter 5.3.5. In chapter 5.3.5, the time afford DT_1 to determine a hearing threshold using the developed procedure is identified.

Individual hearing threshold variations

In Figure 5.10, the individual hearing thresholds L_T measured six times on three (not mandatory consecutive) days are presented. All thresholds are normalized to participant's first measurement.

Hearing thresholds averaged over all participants show no significant differences between measurements⁴, whereat individual thresholds may vary from day to day. A clear tendency towards one preference is however not identifiable.

 $^{^4\}mathrm{ANOVA}$ with repeated measurement: morning to evening for each day seperatly calculated, p < 0.05



Abbildung 5.10: Individual hearing thresholds L_T measured six times on three days. Thresholds are normalized to participant's first measurement.

Aspects of threshold differences and used nomenclature

Level difference $\Delta L_T(f, d, p)$ of a frequency f on a specific day d for a participant p between evening and morning hearing threshold measurement was calculated by:

$$\Delta L_T(f, d, p) = L_T(f, d, p, evening) - L_T(f, d, p, morning)$$
(5.3)

When applying measurement protocol HTM1, L_T is the mean of three separate threshold determinations per frequency and thus the "evening" threshold is given by

$$L_T(f, d, p, evening) = \frac{1}{3} \sum_{i=1}^{3} L_{T_i}(f, d, p, evening), \qquad (5.4)$$

the "morning" threshold by

$$L_T(f, d, p, morning) = \frac{1}{3} \sum_{i=1}^{3} L_{T_i}(f, d, p, morning), \qquad (5.5)$$

where $L_{T_i}(f, d, p, tod)$ is the threshold of one single threshold measurement. Threshold differences $\Delta L_T(f, d, p)$ depend on three variables: frequency f, day d and participant p. Due to the number of frequencies $n_f = 6$, days $n_d = 3$ and participants $n_p = 8$, not all 144 threshold differences are presented in the following, but thresholds are averaged over different variables to characterize dependencies.

Like a TTS, a distinctive threshold difference decreases as the time lapse after exposure increases, which makes it necessary to additionally indicate at which instant after exposure the hearing threshold is measured. For the test in an environment of about 50 dB(A), causing no TTS, there is no need to define a specific time of measurement after exposure. Therefore, the threshold differences are noted as $\Delta L_T(50 \text{ dB}(A))$. However, if time period between end of exposure and point of threshold determination (time of regeneration or restitution) is of relevance, following extended notation is used:

- $\Delta L_{T,2min}$ for a measurement 2 minutes,
- $\Delta L_{T,15min}$ for a measurement 15 minutes and
- $\Delta L_{T,30min}$ for a measurement 30 minutes after exposure.

In addition, the presented noise exposure level is given in brakets, if applicable e.g., $\Delta L_{T,30min}(80 \text{ dB}(\text{A}))$. Averaged mean values are given in $\langle \rangle$ as follows: $\langle \Delta L_{T,30min}(80 \text{ dB}(\text{A})) \rangle$.

Threshold differences as a function of frequency

To investigate whether daily threshold differences depend on frequency, threshold differences are averaged over days to obtain $\Delta L_T(f,p) (= \langle \Delta L_T(f,d,p) \rangle_d)^5$. In a next step the additional average over participants will yield the dependency of daily hearing threshold differences $\Delta L_T(f) (= \langle \Delta L_T(f,d,p) \rangle_{d,p})^6$ only on frequency. Individual frequency dependency $\Delta L_T(f,p)$ are plotted as thin lines and the overall dependency $\Delta L_T(f)$ is drawn as thick line with triangles together with the (sample) standard deviation in Figure 5.11.

⁵Hearing threshold difference as a function of frequency and participants averaged over days: $\Delta L_T(f,p) = \langle \Delta L_T(f,d,p) \rangle_d = \frac{1}{n_d} \sum_d \Delta L_T(f,d,p)$ with n_d number of days (here $n_d = 3$).

⁶Hearing threshold difference as a function of frequency averaged over days and participants: $\Delta L_T(f) = \langle \Delta L_T(f, d, p) \rangle_{d,p} = \frac{1}{n_p} \sum_p \frac{1}{n_d} \sum_d \Delta L_T(f, d, p)$ with n_d number of days (here $n_d = 3$) and n_p number of participants (here $n_p = 8$).



Abbildung 5.11: Hearing threshold differences $\Delta L_T(f, p)$, calculated by subtracting morning threshold from evening threshold (as defined in chapter 5.3.5), averaged over three days, of individual participants p (thin lines) as a function of test frequencies. The thick black line, connecting triangles, is the mean of participants' daily hearing threshold together with (sample) standard deviation $\sigma_{\Delta L_T}$. Right part: Using the same symbols as in the left part the overall mean (black triangle) with standard deviation as well as the individual means of frequency dependent threshold differences $\Delta L_T(p)$ averaged over all frequencies is plotted vertically together with daily threshold difference ΔL_T , averaged over all frequencies f, all days d and all participants p. The characters in the legend are participants codes.

Threshold differences between morning and evening hearing threshold as defined in chapter 5.3.5 shows positive and negative values which means that after work day a participant may be more or less sensitive compared to his/her status at the beginning of work day. The majority of daily threshold differences lies between -2 dB and +2.5 dB with a certain trend to positive values. The variance of daily threshold differences is slightly positive for 250 Hz and 500 Hz tones, becomes minimal (-0.5 dB) for 1 kHz tone and increases then up to 1 dB at 6.3 kHz. Two participants (G and H) show an off-size difference of $\pm 5 \text{ dB}$ for 4 kHz. Statistical tests showed however no significant difference from zero⁷ for different frequencies and for participants, except for participant ('C').

Individual daily hearing threshold differences $\Delta L_T(p)^8$, drawn on right side of Figure 5.11, averaged over all frequencies, lie essentially in a level interval from 0.5 dB to 1 dB. The overall mean ΔL_T^9 , that is additionally averaged over all participants p shows a small positive value of $\Delta L_T = 0.30 \text{ dB} \pm 0.19 \text{ dB}$.

Summing up on average, the hearing threshold at the end of a work day is with $0.30 \text{ dB} \pm 0.19 \text{ dB}$ a little higher than at its beginning but with $\sigma_{\Delta_{L_T}} = 0.38 \text{ dB}$ (95% confidence intervall) statistically not significant¹⁰.

Threshold differences as a function of day

How do hearing threshold differences vary from day to day? Figure 5.12 shows threshold differences $\Delta L_T(p,d) \left(=\langle \Delta L_T(f,d,p) \rangle_f \right)^{11}$ of each participant p and day d, averaged over all frequencies f. Besides the thin lines giving daily threshold differences for individual participants, the thick line represents the mean threshold differences over all participants $\Delta L_T(d) \left(=\langle \Delta L_T(f,d,p) \rangle_{f,p} \right)^{12}$ together with the standard deviation at three different days. On right side of Figure 5.12 the three-day-averages for daily threshold differences of individuals $\Delta L_T(p)$ $\left(=\langle \Delta L_T(f,d,p) \rangle_{f,d} \right)^{13}$ and their mean $\Delta L_T \left(=\langle \Delta L_T(f,d,p) \rangle_{f,d,p_i} \right)^{14}$ are drawn

⁷one-sample t test: equality to zero, p < 0.05

- ¹¹Hearing threshold difference of one day d and one participant p averaged over frequencies and participants: $\Delta L_T(d, p) = \langle \Delta L_T(f, d, p) \rangle_f = \frac{1}{n_f} \sum_f \Delta L_T(f, d, p)$, where n_f is the number of different frequencies (here $n_f = 6$).
- ¹²Hearing threshold difference of one day d averaged over frequencies and participants: $\Delta L_T(d) = \langle \Delta L_T(f, d, p) \rangle_{f,p} = \frac{1}{n_d} \sum_d \frac{1}{n_f} \sum_f \Delta L_T(f, d, p)$, where n_f is the number of different frequencies (here $n_f = 6$) and n_p number of participants (here $n_p = 8$).
- ¹³Hearing threshold difference of one participant p averaged over frequencies and days: $\Delta L_T(p) = \langle \Delta L_T(f, d, p) \rangle_{f,d} = \frac{1}{n_d} \sum_d \frac{1}{n_f} \sum_f \Delta L_T(f, d, p)$ with n_f number of frequencies (here $n_f = 6$) and n_d number of days (here $n_d = 3$).
- ¹⁴Individual daily hearing threshold differences averaged over days and frequencies: $\Delta L_T(p) = \langle \Delta L_T(f, d, p) \rangle_{f,d} = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \Delta L_T(f, d, p)$ with n_f number of days (here $n_d = 3$ and n_f

⁸Individual daily hearing threshold differences averaged over days and frequencies: $\Delta L_T(p) = \langle \Delta L_T(f, d, p) \rangle_{f,d} = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \Delta L_T(f, d, p)$ with n_d number of days (here $n_d = 3$) and n_f number of frequencies (here $n_f = 6$).

⁹Hearing threshold difference averaged over frequencies and days and participants: $\Delta L_T = \langle \Delta L_T(f, d, p) \rangle_{f,d,p} = \frac{1}{n_p} \sum_p \frac{1}{n_d} \sum_d \frac{1}{n_f} \sum_f \Delta L_T(f, d, p)$ with n_f number of frequencies (here $n_f = 6$), n_d number of days (here $n_d = 3$) and n_p number of participants (here $n_p = 8$). ¹⁰also confirmed by an one-sample t test, equality to zero, p < 0.05

vertically with same symbols used on the left. The averaged sample standard deviation of all 144 daily hearing threshold differences is 1.86 dB.

Although there seems to be a trend that daily threshold differences become smaller on second and third day the differences from day to day are too small to become significant: For all participants, the threshold differences are not significantly different from day to day¹⁵. The averaged value for all participants decreases slightly from 0.8 dB to -0.2 dB for the measurement at 3rd day. For participant H, again an off-size difference is noticeable.

For all participants, the averaged threshold differences decrease slightly from 0.8 dB to -0.2 dB for the measurement at 3rd day and are not significantly different from zero¹⁶.

number of frequencies (here $n_f = 6$).

 $^{^{15}\}mathrm{ANOVA}$ with repeated measurement, p < 0.05

 $^{^{16} \}mathrm{one}\text{-sample}$ t test: equality to zero, p < 0.05



Abbildung 5.12: Left part: Hearing threshold differences, calculated by subtracting morning threshold from evening threshold (as defined in chapter 5.3.5), averaged over six tone frequencies, of individual participants (thin lines) for three days of measurement. The thick line, connecting triangles, is the mean of participants thresholds together with standard deviation. Right part: Using same symbols as in the left part the overall mean (black triangle) with standard deviation as well as individual means of daily threshold differences averaged over all three days are plotted vertically. Characters in the legend are participants codes.

Standard deviation

One method to determine the validity of threshold measurements L_T is to calculate standard deviations σ_T of three single threshold measurements $L_{Ti}(i = 1, 2, 3)$, described in Equations (5.4) and (5.5). σ_T should be small for all participants and should not vary considerably during the measurements at different times (tod) and different frequencies (f). It is given in general by Equation (5.6).

$$\sigma_T = \sigma_T(f, P, d, tod) = \sqrt{\frac{1}{2} \sum_{i=1}^3 \left(L_{T_i}(f, P, d, tod) - L_t(f, P, d, tod) \right)^2},$$
(5.6)

with frequency f, participant P, day d and time of day tod. L_{T_i} is the threshold value determined by one single threshold measurement (Equation (5.2)).

To investigate repeatability, the standard deviation of a certain day $d (= \sigma_{\langle d, P \rangle})^{17}$ and of 'time of day' tod $(= \sigma_{\langle tod, P \rangle})^{18}$, averaged per participant P, are shown in Figure 5.13 for all 8 participants. In addition, averaged values for all participants $(\frac{1}{n_P} \sum_P \sigma_{\langle d, P \rangle}, P = 1, ..., 8$; black line with triangles) and all days $(\frac{1}{n_d} \sum_d \sigma_{\langle d, P \rangle}, d = 1, 2, 3$; rightmost dots) are also included.

 $[\]frac{1^{17}\text{calculated by }\sigma_{\langle d,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{d,P} = \frac{1}{n_f} \sum_f \frac{1}{n_{tod}} \sum_{tod} \sigma_T \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \sigma_T \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \sigma_T \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \sigma_T \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} = \frac{1}{n_f} \sum_f \frac{1}{n_f} \sum_d \sigma_T \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18}\text{calculated by }\sigma_{\langle tod,P\rangle} = \langle \sigma_T(f,P,d,tod) \rangle_{tod,P} \\ \frac{1^{18$



Abbildung 5.13: Standard deviations $\sigma_{\langle d,P \rangle}$ and $\sigma_{\langle tod,P \rangle}$ for all 8 participants.

On left side in Figure 5.13, the day dependent standard deviation $\sigma_{\langle d,P\rangle}$ is shown. For two participants (G and H), values increase from about 1.6 dB on first day to 2.7 dB on third day. For all other participants, inter-day difference is below 0.5 dB. Regarding the standard deviations between morning and evening measurement for participants ($\sigma_{\langle tod,P\rangle}$, middle section of the plot), no significant difference¹⁹ was found. On average, the sample standard deviation is 1.41 dB.

Figure 5.14 shows the detailed standard deviations for 3 days ($\sigma_{\langle d,P \rangle}$) with one measurement in the morning and one in the evening ($\sigma_{\langle tod,P \rangle}$).

¹⁹ANOVA with repeated measurement, p < 0.05



Abbildung 5.14: Standard deviations σ for all participants, shown for different days $(\sigma_{\langle d,P \rangle} \text{ with } d = 1, 2, 3)$ and time of days $(\sigma_{\langle tod,P \rangle} \text{ with } tod = \text{morning, evening})$, incl. average, sorted chronologically.

On first day, the (sample) standard deviation is on average 1.25 dB, with a slightly increasing tendency. On second day, the (sample) standard deviation is slightly higher than for the first day at about 1.55 dB with a decreasing tendency. On third day, observed (sample) standard deviation is about 1.4 dB. Variation within one participant from the test in the morning to the following test in the evening is (except for participants 'G' and 'H') below 1 dB.

Figure 5.15 shows frequency dependent (sample) standard deviations $\sigma_{\langle f,P\rangle}^{20}$ for all 8 participants. In addition, averaged values for all participants $(\frac{1}{n_P}\sum_P \sigma_{\langle f,P\rangle}, P = 1, ..., 8$; thick line with triangles) and all frequencies $(\frac{1}{n_f}\sum_f \sigma_{\langle f,P\rangle}, f = 250 \text{ Hz}, ..., 6300 \text{ Hz}$; rightmost dots) are also included.

²⁰standard deviations belonging to a certain frequency f, calculated for each participant P: $\sigma_{\langle f,P \rangle} = \langle \sigma_T(f, P, d, tod) \rangle_{f,P} = \frac{1}{n_d} \sum_d \frac{1}{n_{tod}} \sum_{tod} \sigma_T.$



Abbildung 5.15: Standard deviations $\sigma_{\langle f, P \rangle}$ for all frequencies, shown for different participants, incl. average.

For frequencies from 250 Hz to 2 kHz, averaged (sample) standard deviation is slightly decreasing from 1.44 dB to 1.19 dB. At 4 kHz, an increase to 1.73 dB is observed, followed by an decrease again for 6.3 kHz to 1.30 dB. The increase at 4 kHz is mainly caused again by two participants G and H. Statistical tests confirmed²¹ a significant equality between the mean values and the overall mean.

Summary of standard deviation analysis

Averaged (sample) standard deviation is about $\langle \sigma_T(50 \text{ dB}(\text{A})) \rangle = 1.41 \text{ dB}$. No significant de- or increasing tendency was observed, originating from measurements at different days and different time of days for different measurement frequencies, as shown in Table 5.2. The standard deviation of the mean is given by $\sigma_{\Delta_{L_T}} = 0.19 \text{ dB}$.

²¹one-sample t test: equality to overall mean with p = 0.998)

Cat	day time of			of day	frequency						
σ	σ	σ	σ	$\sigma_{tod=}$	$\sigma_{tod=}$	$\sigma_{f=}$	$\sigma_{f=}$	σ	σ	σ	σ
	d=1	d=3	d=3	morn.	even.	$250~\mathrm{Hz}$	$500 \ Hz$	$1 \mathrm{kHz}$	$2 \mathrm{kHz}$	$4 \mathrm{kHz}$	$6.3 \mathrm{kHz}$
dB	1.25	1.57	1.41	1.41	1.41	1.44	1.49	1.32	1.19	1.73	1.30
dB	mean: 1.41										

Tabelle 5.2: Summary of the observed (sample) standard deviations.

Time requirement of measurement protocols

Another major requirement of the measurement protocol is time consumption required to perform a threshold determination being low. The duration time of one measurement of all 6 frequencies (frequency block FB2, see Table 5.8) is given by $DT_1 = \sum_{f=250 \text{ Hz}}^{f=6.3 \text{ kHz}} DT_{1f}$. In measurement protocol HTM 1, each threshold L_T for a frequency f is measured 3 times ($L_{T1}(f)$ to $L_{T3}(f)$) during one sequence of determination of individual hearing threshold [plus training sequence (measurement of f = 250 Hz, 500 Hz), see FB3 in Figure 5.8]. The average will be used as target value for threshold L_T for frequency f, participant P, day d (in case of repetitions on different days) and time of day tod (morning or evening resp. before or after a noise exposition between two threshold measurements), given by Equation (5.7).

$$L_T = L_T(f, P, d, tod) = \frac{1}{3} \sum_{i=3}^{3} L_{T_i}(f, P, d, tod)$$
(5.7)

On average, participants needed $DT_3 \approx 14 \text{ min } 20 \text{ s}$ to perform a complete set of measurement of hearing threshold tests (as described in Table 5.1). No significant change of duration was observed, neither by day d nor by the time of day tod, shown in Figure 5.16.

Participants showed up slightly different durations of DT_1 each, but variance within one participant at different days was low. Averaged time varies between 2 min 45 s (participant B) to 6 min 4 s (participant H). On average, participants needed 4 min 20 s to perform hearing threshold measurements at 6 frequencies.

In a further analysis, averaged level differences between reversal points were calculated to $5.5 \text{ dB} \pm 0.19 \text{ dB}$, see Appendix for further details.



Abbildung 5.16: Averaged time to perform a hearing threshold measurement at 6 frequencies (DT_1) , incl. mean. For each participant, the time is plotted for the measurement in the morning and in the evening, for all three days, together with an average time. No significant systematical deviation of duration is observed.

5.3.6 Summary of hearing threshold variation without TTS

In a first test campaign, the developed fast audiometric tool was used to test threshold statistics in a non-TTS causing environment. On three (not mandatory consecutive) days, hearing thresholds of eight participants were measured in the morning and in the evening. As a result, an averaged threshold difference $\Delta L_T(50 \text{ dB}(\text{A}))$ within one day of 0.30 dB ± 0.19 dB, averaged over all participants, was observed, which is statistically not significant differences showed up no significant difference. A day dependence of threshold differences with significant differences was also not observed. In addition, an averaged sample standard deviation for threshold determinations σ_T of about $\langle \sigma_T(50 \text{ dB}(\text{A})) \rangle = 1.41 \text{ dB}$ was observed, same holds for different frequencies.

With the developed adaptive measurement system individual hearing thresholds have been determined reproducibly within up to 14 min 20 s (on average). The procedure turned out to be easy in use and did not require additional intervening of the investigator after start.

5.4 Hearing threshold differences after 7 hours at 80 dB(A) exposure

5.4.1 Question

After performing threshold difference tests in a non-TTS causing environment, a second investigation was carried out with an exposure level reported in literature to induce slight temporary threshold difference (see chapter 5.1), detectable two minutes after exposure (defined as $\Delta L_{T,2min}(80 \text{ dB}(\text{A}))$ in chapter 5.3.5). This time, the measuring system, available on six notebooks in parallel, was used in a d = 1 day laboratory experiment to investigate the impact of a 7 hour aircraft cabin interior noise exposure of 80 dB(A) on hearing threshold differences. The presented level corresponds to a cruising condition level of prior generation aircrafts. To gain also information on recovery, additional post-exposure threshold measurements were performed after 15 and 30 minutes (defined as $\Delta L_{T,15min}(80 \text{ dB}(\text{A}))$) and $\Delta L_{T,30min}(80 \text{ dB}(\text{A}))$ in chapter 5.3.5).

5.4.2 Exposure

The laboratory experiment took place in an approximately $50 \square m$ lounge at the University of Oldenburg. Figure 5.17 illustrates the spectra of the sound file, which was presented to the participants by loudspeakers (spectra based on an Airbus A320, which was measured during cruise flight).

5.4.3 Participants

In this experiment, six normal hearing participants (two female and four male, aged between 23 and 29 years) took part.



Abbildung 5.17: Third-octave spectra of sound file, presented 7 hours during the lab test (based on an Airbus A320 - spectra).

5.4.4 Procedure

Just before exposition in the lounge started, individual hearing thresholds $(L_T(f, P, tod = morning))$, see Equation (5.5)), were determined from participants using six measuring stations in parallel in an anechoic chamber, using the measurement system (configuration HTM1). During the 7 hour stay in the lounge, participants were catered and their arrangement of the stay in the room was not constricted, only usage of headphones was not permitted. Immediately after the exposure, further measurements of hearing thresholds (restitution measurements) were performed in the anechoic chamber after 2 minutes, 15 minutes and 30 minutes. Calculating the particular difference between these measurements to the morning measurements $(L_T(f, P, tod = morning))$, this will result in threshold differences, noted as $\Delta L_{T,2min}(80 \text{ dB}(\text{A}))$, $\Delta L_{T,15min}(80 \text{ dB}(\text{A}))$ and

 $\Delta L_{T,30min}(80 \text{ dB}(\text{A}))$ (see chapter 5.3.5 for details). Differing from the test in the morning, now the six test frequencies were measured only once (protocol HTM2). Figure 5.18 shows again a schematic plot of timeline.



Abbildung 5.18: Schematic plot of timeline during the 7 hour / 80 dB(A) exposure. In the beginning, hearing thresholds were measured using measurement protocol HTM1. After exposure, hearing thresholds were measured again, using protocol HTM2, after 2, 15 and 30 minutes to determine $\Delta L_{T,2min}(80 \text{ dB}(A)), \Delta L_{T,15min}(80 \text{ dB}(A))$ and $\Delta L_{T,30min}(80 \text{ dB}(A))$, as defined in chapter 5.3.5.

5.4.5 Results

Threshold differences

Figure 5.19 shows differences of hearing thresholds measurement afterwards $(L_{tod=evening} = \Delta L_{T,2min}, \Delta L_{T,15min}, \Delta L_{T,30min})$ for each participant in comparison to the measurement before exposition $(L_{tod=morning})$, averaged over all frequencies²². In addition, the averaged threshold differences for all participants²³ are shown (thick line connecting triangles), together with the (sample) standard deviation $\langle \sigma_{\Delta L_T(2,15,30),P} \rangle$ given by

$$\left\langle \sigma_{\Delta L_T(2,15,30),P} \right\rangle = \sqrt{\frac{1}{n_P - 1} \sum_P \left(\Delta L_{TP} - \Delta L_T \right)^2} \tag{5.8}$$



Abbildung 5.19: Temporary hearing threshold differences for six participants 1 to 6 after 2, 15 and 30 minutes in comparison to determined thresholds before exposition (thin lines, averaged over all frequencies) and averaged over all participants (thick line connecting triangles).

All observed hearing threshold differences decrease with time after exposition monotonously (indicating a recovery of hearing threshold towards the level of preexposition), except for participant '5'. Statistically, a significant threshold difference was observed²⁴ for five of six participants at the measurement 2 min after exposure $(\Delta L_{T,2min}).$

Threshold differences as a function of frequency

Differences of hearing thresholds measurement 2, 15 and 30 minutes afterwards $(L_{tod=evening} = \Delta L_{T,2min}, \Delta L_{T,15min}, \Delta L_{T,30min})$ for each frequency in comparison to the measurement before exposition $(L_{tod=morning})$, averaged over all participants²⁵

²²Threshold differences averaged over frequencies f are calculated by $\langle \Delta L_{T(2,15,30),P} \rangle =$ $\frac{\frac{1}{n_f} \sum_f \Delta L_{T2,15,30}(f, P)}{\text{calculated by } \frac{1}{n_f} \sum_f \frac{1}{n_P} \sum_P \Delta L_{T2,15,30}(f, P)}$

 24 ANOVA with represented measurement, p < 0.05

²⁵Threshold differences averaged over participants P are calculated by $\langle \Delta L_{T(2,15,30),f} \rangle =$ $\frac{1}{n_P} \sum_P \Delta L_{T2,15,30}(f, P).$

are shown in Figure 5.20.

In addition, the averaged threshold differences for all frequencies²⁶ are shown (thick line connecting triangles), together with the (sample) standard deviation $\langle \sigma_{\Delta L_T(2,15,30),f} \rangle$ given by the adaption of Equation (5.8).



Temporary threshold differences ΔL_{τ}

Abbildung 5.20: Temporary hearing threshold differences for 6 measured frequencies after 2, 15 and 30 minutes in comparison to determined thresholds before exposition, averaged over all participants and frequencies (thick line, same as in Figure 5.19). The negative value for 4000 Hz at $\Delta L_{T,30min}$ is most likely explained by the measuring accuracy.

2 minutes after a 7-hour exposure to 80 dB(A) cabin noise, a statistically significant averaged threshold difference of 3.9 dB \pm 0.51 dB was determined. 15 minutes after exposure, a difference of 1.7 dB \pm 0.50 dB, again statistically significant, was observed. After 30 minutes, a 0.3 dB \pm 0.48 dB difference was found (not statistically significant). Each difference was tested with an ANOVA with repeated measurements, p < 0.05.

²⁶calculated by $\frac{1}{n_f} \sum_f \frac{1}{n_P} \sum_P \Delta L_{T2,15}_{,30}(f,P)$

Analysis of (sample) standard deviation

In addition, an analysis of differences in standard deviation of the measurement in the morning (using measurement protocol HTM1^{27}) was performed. Therefore, all (sample) standard deviations belonging to a certain frequency f are calculated for each participant P:

$$\sigma_{\langle f,P\rangle} = \langle S_T(f,P,d=1,tod=morning) \rangle_{f,P} = S_T$$
(5.9)

Furthermore, averaged values for all participants $(\frac{1}{n_P}\sum_P \sigma_{\langle f,P \rangle}, P = 1, \dots, 8;$ thick line with triangles) and all frequencies $\frac{1}{n_f}\sum_f \sigma_{\langle f,P \rangle}, f = 250$ Hz, ..., 6300 Hz, rightmost) are also included in Figure 5.21.



Abbildung 5.21: Standard deviations (sample) for all frequencies, shown for all 6 participants, incl. average $\langle \sigma_T (80 \text{ dB}(\text{A})) \rangle_{morning} = 1.27 \text{ dB}.$

From low frequencies to high frequencies, the averaged (sample) standard deviation of the measurement in the morning (HTM1) is slightly decreasing from 1.6 dB at 250 Hz to 0.9 dB at 6300 Hz. The overall average (sample) standard deviation is calculated to $\langle \sigma_T (80 \text{ dB}(\text{A})) \rangle = 1.27 \text{ dB}$. Statistical tests²⁸ showed a significant equality between the mean values of all frequencies and the overall mean.

²⁷In HTM2, used in the evening, no repeated measurements are performed.

²⁸one-sample t test: equality to $\langle \sigma_T (80 \text{ dB}(\text{A})) \rangle = 1.27 \text{ dB}, p < 0.05$

5.4.6 Summary of hearing threshold variation in TTS-causing environment

In a second acoustic environment, the developed audiometric tool was used during a 7 hour lasting experiment with a TTS causing sound level. Two minutes after exposure with 80 dB(A) cabin interior noise, this investigation revealed a (statistically significant) temporary threshold difference in hearing of about 3.9 dB±0.51 dB on average. As a function of time, the observed threshold difference decreases after 15 minutes to 1.7 dB ± 0.50 dB (also significant) and to 0.3 dB ± 0.48 dB (not significant) 30 minutes after end of exposure. The threshold difference for the measurement 30 min after ($\Delta L_{T,30min}$) is comparable to the result of the 50 dB(A) test (0.3 dB ± 0.19 dB, see chapter 5.3.6).

Average (sample) standard deviation is calculated to $\langle \sigma_T (80 \text{ dB}(\text{A})) \rangle = 1.27 \text{ dB}$ and therefore lower than for the 50 dB(A) test with $\langle \sigma_T (50 \text{ dB}) \rangle = 1.41 \text{ dB}$, as given in Table 5.2.

In this experiment, 30 min after a 7 hours lasting 80 dB(A) noise exposure, hearing thresholds had recovered and no TTS could be observed any longer.

5.5 Threshold measurements before and after simulated flights

5.5.1 Question

To investigate potential temporary threshold differences ΔL_T in aircraft cabins during long-haul flights, individual hearing thresholds of 108 participants in total were measured within the ICE-project before and after simulated long haul flights (7 hours) in order to verify a possible occurrence of a threshold difference. During these flights, three in level different realistic noise conditions (64 dB(A), 69 dB(A) and 74 dB(A), one per day) were presented in a ground-based cabin simulator.

5.5.2 Exposure

A third-octave spectra of the exposure, which is based on a recording performed during a cruise level fight with an Airbus A320, is shown in Figure 5.22.



Abbildung 5.22: Third-octave spectra (74 dB(A), exemplary for 6th row), presented during a simulated 7-hour flight inside an Airbus A310 simulator. Slight deviations may occur from seat to seat.

5.5.3 Participants

For measurements of individual hearing thresholds before and after 7 hour exposure, up to six passengers per day were available; tests were performed on 18 days. 108 different passengers in total participated in threshold measurement, subdivided into 84 passengers at 74 dB(A), 6 at 69 dB(A) and 18 in 64 dB(A) condition. All participants were at one's own opinion not limited in normal hearing. An age distribution is given in the Appendix, Figure 5.29.
5.5.4 Procedure

After measuring pre-flight individual hearing thresholds right before "take-off" (L_{pre}) , the participants joined together with the remaining 34 passengers of the day's investigation the simulated flight at allocated seats in aircraft cabin.

Due to operational limits within the ICE-project, the second measurement of individual thresholds post-flight (L_{post}) could at the earliest be performed 30 minutes after end of noise exposure, enabling us to determine the $\Delta L_{T,30min}$. A schematic plot is given in Figure 5.23. Threshold tests were performed in an office room where damping capabilities of closed headphones were sufficient to prevent volunteers from being disturbed by background noise during threshold measurements. For both, pre-flight and post-flight measurements, the measurement protocol HTM1, shown in Table 5.1, has been used. Having answered all open questions from volunteers by the investigator, all 6 participants performed the test in parallel at six stand-alone measuring stations.



Abbildung 5.23: Schematic plot of timeline during the tests within the ICE-project. Before exposure, hearing threshold measurements using protocol HTM1 were performed. During the following 7 hours the noise in the cabin was adjusted to levels of either 64 dB(A), 69 dB(A) or 74 dB(A) (depending on the test plan) in the course of a simulated flight. 30 min after exposure end, hearing thresholds were measured again using HTM1 protocol. By subtracting, $\Delta L_{T,30min}$ is calculated, as described in chapter 5.3.5.

5.5.5 Results

A first check of the data validity showed that the results of 20 out of 108 participants had to be excluded for further analysis (by reason of lack of co-operation (5 participants), early experiment break off (6) and conductive deafness due to a pressurized inner-ear (9)).

Threshold differences

Again hearing threshold differences between post- and pre-flight served as an indicator for noise effects. For each of the remaining 88 participants, the averaged threshold difference is plotted in Figure 5.24 above particular sound pressure level, the participant was exposed to during the flight at its seat. Sound field distribution was identified in detail during an earlier pre-experiment of the ICE-project. In addition, threshold differences averaged over all participants of each condition are plotted (black line). Also, the (sample) standard deviation is given as error bar for each participant.

For 64 dB(A)-condition, the averaged threshold difference 30 min after exposure to morning measurement was determined²⁹ for 12 participants and resulted to 0.7 dB \pm 0.24 dB, which is significantly different from zero (based on a one-sample t test, equality to zero, p < 0.05). 6 participants were exposed to the 69 dB(A) flight condition and showed up a threshold difference of 0.6 dB \pm 0.50 dB (statistically not significant different from zero). In 74 dB(A) condition, 70 participants took part and a temporary threshold difference of 0.7 dB \pm 0.12 dB was determined on average (again significantly different from zero).

Statistical tests³⁰ confirmed a significant threshold difference for one participant after 64 dB(A)-condition (8.3%, N = 12), two participants after 69 dB(A)-condition (33%, N = 6) and 15 participants after 74 dB(A)-condition (reduced hearing performance observed for 13 participants, 18.6%, N = 70, and improved performance for two participants, 2.9%, N = 70).

 $^{^{29}}$ as described in chapter 5.3.5

 $^{^{30}}$ one-sample t test: equality to zero, p < 0.05



Abbildung 5.24: Differences in hearing thresholds ($\Delta L_{T,30min}$, subtracting 'after flight' from 'before flight' measurement, as defined in chapter 5.3.5) for 88 passengers are plotted above sound pressure level, each participant was exposed to during the flight at corresponding seat (sound field distribution was identified in detail during an earlier pre-experiment of the ICE-project). (Sample) standard deviation is given as error bar for each participant. In addition, the differences averaged over all passengers in each condition (black triangles) are connected with a line for visualisation.

Analysis of (sample) standard deviation

For validation, the (sample) standard deviation $\sigma_T(f, P, tod)$ of measurements were calculated (5.10), separately for all participants for morning and evening measurement.

$$\sigma_T = \sigma_T(f, P, d, tod) = \sqrt{\frac{1}{2} \sum_{i=1}^3 \left(L_{T_i}(f, P, d, tod) - L_t(f, P, d, tod) \right)^2}$$
(5.10)

Next, all (sample) standard deviations belonging to a certain frequency f are calculated for those six participants P(d) (with pax = P(d)), taking part in one

days' simulated flight (= day = d with one flight per day), separately for time of day (tod = morning, evening):

$$\sigma_{\langle f, pax \rangle} = \langle \sigma_T \left(f, pax, tod \right) \rangle_{f, pax} = \frac{1}{n_{pax}} \sum_{pax} \sigma_T \tag{5.11}$$

In Figure 5.25, also the average values over all flights (black line) for each frequency³¹ and the average over all frequencies for tod = morning, evening (separate dots aside)³² and both (rightmost dots)³³ are given.

The observed averaged (sample) standard deviation for all conditions in the morning is $\sigma_{tod=morning} = \langle \sigma_T(f, pax, d, tod) \rangle_{tod=morning} = 1.53 \text{ dB}$ and in the evening $\sigma_{tod=evening} = \langle \sigma_T(f, pax, d, tod) \rangle_{tod=evening} = 1.31 \text{ dB}$ (decreasing from 1.48 dB to 1.05 dB). The combined average (sample) standard deviation is given by $\sigma_T = \langle \sigma_T(f, pax, d, tod) \rangle = 1.42 \text{ dB}$. Based on one-sample t test³⁴, 1 kHz and 2 kHz for morning and 4 kHz as well as 6.3 kHz for evening measurement are significantly different from overall mean.

³¹Frequency dependent standard deviation averaged over all flights f derived by $\sigma_{\langle f \rangle} = \langle \sigma_T(f, pax, d, tod) \rangle_f = \frac{1}{n_d} \sum_d \frac{1}{n_{pax}} \sum_{pax} \sigma_T$ ³²Frequency dependent standard deviation averaged over different time of day *tod* derived by

³²Frequency dependent standard deviation averaged over different time of day tod derived by $\sigma_{\langle tod \rangle} = \langle \sigma_T (f, pax, d, tod) \rangle_{tod} = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \frac{1}{n_{pax}} \sum_{pax} \sigma_T$ ³³Frequency dependent standard deviation averaged over all flights f and different time of day tod

³³Frequency dependent standard deviation averaged over all flights f and different time of day tod derived by $\sigma_T = \langle \sigma_T (f, pax, d, tod) \rangle = \frac{1}{n_f} \sum_f \frac{1}{n_d} \sum_d \frac{1}{n_{pax}} \sum_{pax} \sigma_T$

³⁴testing equality to averaged σ_T



Abbildung 5.25: Standard deviations (sample) for the participants in one day's flight, separately derived for the measurement in the morning (left) and in the evening (right), each with averaged values (for all flights (black line), all morning, all evening (single dots)) as well as the overall average (single dots rightmost).

5.5.6 Summary of hearing threshold variations before and after simulated flights

To answer the question whether specific noise in an aircraft cabin may cause a temporary physiological impairment of hearing organ in terms of temporary threshold shifts, individual hearing thresholds of 108 participants in total were measured within the ICE-project before and after simulated flights with three different noise conditions, one per day. For analysis, results of 88 participants were used.

Observed threshold differences (difference between 30 min after exposure to measurement in the morning) for presented noise conditions of 64 dB(A), 69 dB(A) and 74 dB(A) are listed in Table 5.3, 2nd row. Significant differences, based on one-sample t test (equality to zero, p < 0.05), are marked with (*). Some participants developed a significant positive threshold difference³⁵ between evening and morning measurement, as shown in Table 5.3, 4th row. For 74 dB(A), also two participants showed a significant hearing improvement, values given in brackets.

Tabelle 5.3: For each condition (described in 1st column), the averaged observed threshold difference $\Delta L_{T,30min}$ is given in 2nd column. The number of participants performing the test (3rd column) and the amount of passengers with an observed significant threshold differences (decrease of hearing capability, 4th) are following. Last column gives the relative number in %. For 74 dB(A), also two participants showed a significant hearing improvement, values given in brackets.

	observed	number no. of participants			
condition	threshold	of	with significant	proportionally	
	difference	participants	threshold differences		
$64 \mathrm{dB}(\mathrm{A})$	0.7 dB *	12	1	8.3%	
$69 \mathrm{dB}(\mathrm{A})$	$0.6~\mathrm{dB}$	6	2	33%	
74 dB(A)	$0.7~\mathrm{dB}$ *	70	13(2)	18.6%~(2.9%)	

As a result, the threshold difference 30 min after exposure for the 64 dB(A) condition is in range of the difference, observed for the 74 dB(A) condition (both show a significant difference from zero, based on a one-sample t test (p < 0.05). The determination of threshold differences for the 69 dB(A)-condition is based on six participants (result not significant different from zero).

The averaged (sample) standard deviation is calculated to $\sigma_T = 1.42 \text{ dB}$ and therefore comparable to the observed (sample) standard deviation of 50 dB(A) tests $(\langle \sigma_T (50 \text{ dB}(A)) \rangle = 1.41 \text{ dB}).$

In another analysis, also reversal point differences were calculated and resulted in a $\Delta L_{rev} = 4.95 \text{ dB}(\text{A}) \pm 0.04 \text{ dB}(\text{A})$, see Appendix for details.

5.6 Summary of results

To give an overview, threshold differences of all experiments performed are listed in the Table 5.4 (incl. standard deviation). Significant differences (from zero), based

³⁵based on one-sample t test: equality to zero, p < 0.05.

on one-sample t test, are marked (*). For all tests, a positive threshold difference was observed.

Tabelle 5.4: Observed threshold differences for 2, 15 and 30 minutes after exposure $(3^{rd} \text{ to } 5^{th} \text{ column})$ and standard deviations (in brakets) of all experiments performed (noise exposure level shown in 1^{st} column). 2^{nd} column gives the number of participants. Results are **sorted by level** and not in order of the experiments performed. Significant threshold differences (from zero, p < 0.05) are marked (*).

noise	number	observed difference ΔL_T			
exposure	of	(standard deviation)			
during test	participants	$2 \min$	$15 \min$	$30 \min$	
50 dB(A)	8	0.3 dB (0.19 dB)			
$64 ext{ dB}(A)$	12	_	_	$0.7 \text{ dB}^* (0.24 \text{ dB})$	
$69 \mathrm{dB}(\mathrm{A})$	6	—	—	0.6 dB (0.50 dB)	
74 dB(A)	70	—	—	$0.7 \text{ dB}^* (0.12 \text{ dB})$	
80 dB(A)	6	3.9 dB*	1.7 dB^*	0.3 dB	
		(0.51 dB)	(0.50 dB)	(0.48 dB)	

Significant threshold differences³⁶ were found 2 min and 15 min after an exposure of 80 dB(A). Also, 30 min after an exposure of 64 dB(A) and 74 dB(A) significant threshold differences were observed. All other differences are noch significantly different from zero.

5.7 Conclusion

Within the EU research project "Ideal Cabin Environment" (ICE) individual hearing thresholds of 108 participants were determined before and after simulated 7 hour long-haul flights to answer the question, if noise exposure at a level of 74 dB(A) can cause physiological impairment of auditory system like a time limited (temporary) threshold shift (TTS). To measure threshold differences ΔL_T of air travel passengers, a mobile, fast and accurate audiometric tool was developed, tested and used to determine the statistics of normal hearing threshold variations. For an exposure of 50 dB(A), where no TTS is expected, an averaged threshold difference within one

 $^{^{36}\}mathrm{based}$ on an ANOVA with repeated measurements / one-sample t test, equality to zero, p < 0.05

day of $0.30 \text{ dB} \pm 0.19 \text{ dB}$, averaged over all participants, was observed, which is statistically not significant different from zero (t test, p < 0.05).

In a second experiment, noise (spectrally shaped as aircraft cabin interior noise) of a level of 80 dB(A) was presented to six participants during 7 hours. Threshold shifts two minutes after exposure were determined as well as the restitution curves (further measurements after 15 min and 30 min. After 2 min, the observed threshold difference of $3.9 \text{ dB} \pm 0.51 \text{ dB}$ and after 15 min the difference of $1.7 \text{ dB} \pm 0.50 \text{ dB}$ are significantly different from zero and therefore identified as TTS, whereas after 30 min (0.3 dB ± 0.48 dB) no significant difference was found.

Results of threshold difference measurements 30 min after noise exposure for 7 hours at a level of 74 dB(A) in an aircraft simulator lead to a difference of 0.7 dB \pm 0.12 dB, which is significantly different from zero and therefore identified as a threshold shift of small magnitude.

5.8 Acknowledgement

The authors are very grateful to Berit Jütz and Nils Freese for their support in performing the measurements and Ping Rong for coding the measurement routine. We would also like to thank all volunteers participating in the different studies as well as Dr. Roland Kruse and Daniela Poppinga for assistance with the statistical analysis.

The European commission financially supported the ICE project under contract number AST4-CT-2005- 516131. The authors wish to thank the consortium consisting of the following partner organisations:

- BRE, United Kingdom, Project Coordinator
- Airbus Deutschland GmbH, Germany
- Antanas Gustaitis Aviation Institute of Vilnius Gediminas Technical University, Lithuania
- Avitronics Research, Greece
- Civil Aviation Authority, Aviation Health Unit, United Kingdom
- Deutsches Zentrum für Luft- und Raumfahrt e. V., Germany
- EADS Innovation Works, Germany
- Fraunhofer Institut für Bauphysik, Germany
- Medical University of Vienna, Austria
- Stichting National Lucht en Ruimtevaartlaboratorium, The Netherlands
- Royal Free and University College Medical School, United Kingdom
- Streit Technische Gebäude Ausrüstung, Germany
- Carl-von-Ossietzky Universität Oldenburg, Germany
- Universita delgi Studi di Padova, Italy
- Czech Technical University in Prague, Czech Republic

5.9 Appendix

Instruction to participants



Abbildung 5.26: One of the instruction to the participant as an example, shown on the notebook display during hearing threshold measurement ("Please press any key, once the tone is audible."). In addition, there is a pictogram shown, which indicates that the level of test signal starts below threshold. The percentage value at the bottom gives information on progress.

Analysis of reversal points

For calculation of hearing threshold L_T the mean of last four reversal points is employed: $L_T = \frac{1}{4} \sum_{k=5}^{8} L_{rev_k}$. Level difference between upper and lower reversal points may be due to individual differences that relate to certainty of detecting presence or absence of a tone. The level difference ΔL_{rev} between two adjacent reversal points is given by $\Delta L_{rev}(n) = |L_{rev_{n+1}} - L_{rev_n}|$. It will be investigated regarding possible parameters as participant p, frequency f, day d and time of day tod. The threshold is determined as mean of four last reversal points. To characterize level differences between upper and lower reversal points the mean L_{rev} of last four reversal point differences are chosen, based on Equation (5.12).

$$L_{rev} = \frac{1}{4} \sum_{n=4}^{7} L_{rev}(n)$$
(5.12)

Of course, L_{rev} may depend on participant P, frequency f, day d and time of day tod: $L_{rev} = L_{rev}(P, f, d, tod)$ and dependencies of different parameters will be shown

by averaging over the others.

Test in non-TTS environment

In Figure 5.27, the participant's P averaged level differences $\Delta L_{rev}(d, tod)^{-37}$, $\Delta L_{rev}(tod)^{38}$, $\Delta L_{rev}(d)^{39}$ and $\Delta L_{rev}(f, d)^{40}$ between reversal points (see Figure 5.4 for details) for each frequency f, based on Equation (5.12), were plotted. In addition, averaged values for all participants $(\frac{1}{n_F}\sum_{P}\Delta L_{rev\langle f,P\rangle}, P = 1, \ldots, 8,$ black line with triangles) and all frequencies $(\frac{1}{n_f}\sum_{f}\Delta L_{rev\langle f,P\rangle}, rightmost dots)$ are also included.

For all frequencies, average level difference between reversal points is about 5.5 dB. Statistical tests⁴¹ showed no significant difference from frequency to frequency.

The shown (sample) standard deviations (whiskers) for each participant are calculated by

$$\sigma_{\Delta L_{rev}} = \sigma_{\Delta L_{rev}}(f, P, d, tod) = \sqrt{\frac{1}{2} \sum_{d=1}^{3} \left(\Delta L_{rev_d} - \left\langle \Delta L_{rev}(f, P, d, tod) \right\rangle \right)^2} \quad (5.13)$$

Participants showed up a frequency-independent level difference between reversal points of about $\Delta L_{rev} = 5.5$ dB on average. Between measurements on 3 different days, participants only show small variations ($max (\sigma_{\Delta L_{rev}}) = 0.75$ dB, on average $\langle \sigma_{\Delta L_{rev}} \rangle_{f,P,tod,d} = 0.49$ dB. The standard deviation of the mean is given by 0.19 dB.

Test in simulated flight environment

Again, the over all participants of the simulated flight environment test averaged level differences ΔL_{rev} between the reversal points (see Figure 5.27 for details, procedure described in chapter 5.9, e.g., Equation (5.12) ff) for each frequency f was



Abbildung 5.27: Averaged level differences ΔL_{rev} between reversal points for each participant of the non-TTS environment test, averged over frequency f as well as averged over p are shown.

plotted in Figure 5.28.

Level differences between reversal points are nearly constant with only a slight decrease from 250 Hz (5.11 dB) to 6300 Hz with 4.70 dB. On average, a level difference of $4.95 \text{ dB} \pm 0.04 \text{ dB}$ (averaged sample standard deviation 0.91 dB) was observed.

37Level difference averaged over days d and 'time of day' tod: $\langle \Delta L_{rev}(tod,d) \rangle_{f,p}$ = $\begin{array}{c} \frac{1}{n_d}\sum_{d}\frac{1}{n_{tod}}\sum_{tod}\Delta L_{rev}(tod,d)\\ ^{38}\text{Level} \quad \text{difference} \quad \text{averaged} \end{array}$

'time $\langle \Delta L_{rev}(tod) \rangle_{f,P}$ of day' tod:over $\tfrac{1}{2}\left(\Delta L_{rev}('morning') + \Delta L_{rev}('evening')\right)$

³⁹Level difference averaged over days $d: \langle \Delta L_{rev}(d) \rangle_{f,P} = \frac{1}{n_d} \sum_d \Delta L_{rev}(d)$ ⁴⁰Level difference averaged over frequencies f and days $d: \langle \Delta L_{rev}(d, f) \rangle_{f,P}$ $\frac{1}{n_f}\sum_f \frac{1}{n_d}\sum_d \Delta L_{rev}(d,f)$ 41 one-sample t test: equality to mean, p<0.05



Abbildung 5.28: Averaged level differences ΔL_{rev} between the reversal points, averaged over all 88 participants of simulator test. Error bars show sample standard deviation.

These level differences are comparable to the step width of 5 dB, used for standard audiometry as defined in ISO/DIS/8253.2 [70]. In comparison to the developed method, standard audiometry requires one investigator per participant to operate the audiometer. Though in diagnostic medicine it is a standard procedure, it is not applicable for field tests with a large amount of participants to be investigated in parallel (as defined in the requirements).

Age distribution



Abbildung 5.29: Age distribution of all passengers participating in hearing threshold investigation before and after a simulated 7-hour flight. Age distribution of all passengers in ICE experiment was equally distributed. For hearing threshold tests, about 50% of participants were chosen to be aged 18 to 28.

6 Impact of sound and vibration on health, travel comfort and performance of flight attendants and pilots.

Im EU-Projekt ICE stand die Wirkung von Umgebungsparametern auf den Komfort, das Wohlbefinden und den allgemeinen Gesundheitszustand von Passagieren im Vordergrund (siehe Kapitel 2). Jedoch dient die Flugzeugkabine nicht nur als Beförderungsmittel, sondern ist ebenfalls der Arbeitsplatz von Flugbegleitern und Piloten. Die Wirkung der physikalischen Umgebungsparameter, insbesondere Schall und Vibrationen, auf Gesundheit, Komfort und Arbeitsleistung der Crew am Arbeitsplatz sind im Rahmen des Forschungsprojekts HEACE in simulierten und realen Flügen untersucht worden. Die hier dargestellten Ergebnisse wurden vom Elsevier-Verlag¹ veröffentlicht in:

Aerospace Science and Technology, Volume 12, Issue 1, January 2008, Pages 18-25 Autoren: Volker Mellert, Ingo Baumann, Nils Freese, Reinhard Weber. Eine weiterführende Diskussion erfolgt in Kapitel 7 ab Seite 144.

6.1 Abstract

Noise and vibration affect flight attendants and pilots at their workplace in cabin and cockpit, respectively, besides numerous other physical environmental parameters, e.g. air quality, draft, temperature distribution, glare, etc.. Additionally, the physiological setting and psychological attitude modulate the impact of environmental parameters. The relation between environmental parameters at the specific work place and the subjective and medical response and performance was investigated in the EU-project HEACE (www.heace.org). Investigations were conducted in simulators and mock-ups as well as in real long-haul flights. Environmental measurements were performed with the help of the partners EADS-CRC, BRE, itap GmbH and Paragon Ltd. Physiological and medical data were recorded by the Medical University Vienna. Questionnaires were designed with support of BRE. A comprehensive data

¹http://ees.elsevier.com/aescte, ISSN: 1270-9638, Imprint: Elsevier

analysis was performed with BRE, CIRA and University of Patras. The sound level is one important parameter besides air quality with negative impact on subjective feeling and significant bias on the crew's performance, in particular when effective during long periods.

6.2 Introduction

The investigation of environmental impact in the aircraft cabin on performance, well-being and health of cabin- and flight-crew is the objective of the European project HEACE ("Health Effects in Aircraft Cabin Environment", www.heace.org). One aim of the research is to develop a model which relates the impact of the ambience of the workplace of the crew to health parameters, subjective comfort, performance, well-being etc. of the crew. The model should help to derive recommendations for an improved design of this workplace. Additionally, investigations are conducted in collaboration with the European technology platform FACE, where passengers' ride comfort is investigated with a similar objective (to derive a so-called passenger comfort index) [7].

HEACE is carried out together with partners from Building Research Establishment (BRE) in UK, Medical University Vienna in Austria, EADS-CRC in Germany, CIRA in Italy, University of Patras (LFME) in Greece, itap GmbH in Germany, and Paragon Ltd in Greece.

6.3 Test design

6.3.1 Premise

Variation of environmental conditions in real flight is obviously limited, and experimental set-ups in an airplane have to follow rigid boundary conditions by taking safety issues into account. It is therefore useful to conduct experiments in a simulator facility which provides a sufficient natural ambience with the help of virtual reality and an appropriate mock-up environment in hardware. Experiments are carried out in the HEACE project in both simulators and during real flights. The question is still unsolved how "real" a simulation has to be conducted in order to allow transfer of results from the test-bed to the real-flight situation. The HEACE project could not answer this question, but is the first investigation, which uses systematically in parallel the same methodology in both, real flight and mock-up experiments. The environment in real flights cannot (really) be changed by purpose, but monitored with data logging. Since the environmental condition varies with time and with location within the cabin, a certain spread of parameters is guaranteed. The environmental condition in the simulator is in general more-or-less adjustable according to the specific experimental design.

Flight experiments serve in a first phase of the investigation to set up a data base with physical data of the cabin and cockpit environment in order to feed the simulator facilities with a realistic noise, vibration, and climatic environment. Main flight experiments are conducted in 6 long-haul flights of 8 and 12 hour duration, in which altogether 132 flight attendants and 30 pilots participated. The environmental data are registered as "naturally" offered; the data from the cabin crew is recorded with the same tools and by the same means as used in the simulator experiments under precisely defined conditions. The design comprises a questionnaire with paragraphs to be used before, during and after the test period. The questionnaire employed in the simulation tests is a little bit more detailed and includes e.g. "mood" queries. Additionally, numerous physiological and health indices are registered.

Simulator tests are carried out in two facilities: The emergency trainer of Austrian Airlines in Vienna and in the ACE test facility of BRE in Watford². The environmental condition is adjusted in a full three-step variation of the three parameters "sound and vibration" (vibration signal derived from the acoustic input), "humidity", and "temperature". The parameters are recorded throughout cabin and cockpit, because a certain spatial variation is inevitable, although a target value is adjusted in the simulator for the respective step. 22 pilots and 86 flight attendants serving 544 (test) passengers participated in the simulator tests.

6.3.2 Questionnaire design

Questionnaires are designed based on own extensive experience in psychophysical research for car industry, and on the European project IDEA PACI [71] as well as on the knowledge gained from the European project CabinAir [6]. The questions base on interviews with the respective personnel and on extraction of appropriate pairs of adjectives, differentially describing the cabin and cockpit environment.

²http://projects.bre.co.uk/envdiv/aviation/test.html, July 2009

These semantic differentials are tested with experts. Mood scales, questions related to symptoms and subjective activity are included. Design and layout of appropriate scales (in general a 7-point scale as fixed by other guidelines (e.g. [72])) is formatted for electronically reading. The scales could either be bipolar (e.g. -3 to +3), or sometimes unipolar (e.g. 1-7 or 1-5). After estimating the time to fill-in the questionnaire, the time frame for conducting the experimental procedure in simulators and in real flights is included in the experimental design. In general, the questions related explicitly to the environmental and subjective condition just perceived within the period when the questionnaires are filled in.

Questions address some 120 items from following areas

- health and well-being (30 items)
- environmental conditions (45 items)
- demand for control over environment (8 items)
- effect of the environment (18 items)
- relative comfort contribution (18 items)
- ability to work (8 items)
- alertness and mood (9 items).

Fig. 6.1 gives an impression of the distribution of questionnaires to passengers in the ACE simulator. The members of the cabin crew are "wired" to a data logger, which register physiological indicators.

The questionnaires are designed by the partners from Oldenburg University, Medical University Vienna and BRE. The questionnaire for the passengers (in the FACE project) is similar, except for work-related items, and with additional comfort-related items.

 $6 \,$ HEACE – Impact of sound and vibration on health, travel comfort and performance of crew



Abbildung 6.1: Distribution of questionnaires to "test passengers" in the ACE simulator during a 3-hour simulated flight. The flight attendants carry mobile data loggers to monitor their physiological status.

6.3.3 Flight test

Measurements are made with support of Austrian Airlines during six long-haul flights: Vienna-Delhi-Vienna (8 h duration) and Vienna-Tokyo-Vienna (12 h duration) in Airbus A330 and A340.

Environmental data are measured in cockpit, galleys, cabin, and crew-rest compartments at various locations. The experiments and data acquisition are carried out with the partners of Oldenburg University, Medical University Vienna, EADS-CRC, itap GmbH and Paragon Ltd.

Following environmental parameters are measured:

• sound and vibration (time history and level)

- temperature
- humidity
- draft
- air quality (CO₂, CO, VOCs, number of germs).

Practically coherent in the same time, the following physiological parameters are measured in parallel:

- heart rate and -variability
- blood pressure
- oxygen saturation
- salivary cortisol
- skin conductance

while questionnaires are filled in by the crew members. This is scheduled after service (three-times during Tokyo flight and two-times during Delhi flight).

An example for the environmental condition with respect to sound pressure level is given in Fig. 6.2, which notes the A- and B-levels at different locations of the planes. A difference between A- and B-weighted level indicates an additional low-frequency contribution (most pronounced in the first economy compartment in the middle of the fuselage near the wings). Both planes have about the same noise level in cabin and cockpit during the flight tests except for the aft of the A330, where the level is slightly increased (due to the two louder engines).

Fig. 6.3 gives the dB(A) level from the cockpit location to the aft, averaged over all flights and distinguished between the two types of aircraft (A330 and A340). A continuous increase with location is observed. Fig. 6.4 indicates in particular an increase of low-frequency components starting at the "economy 1"-location. The vibration levels of acceleration are measured at about the same locations and depicted in Fig. 6.5.

The acceleration levels in Fig. 6.5 are averaged values in the frequency bands 16-80 Hz, 100-160 Hz, 200-250 Hz and weighted according to ISO 2631-1. Note that



Abbildung 6.2: Distribution of A- vs. B-weighted sound level in cockpit and cabin for A340 (filled symbols) and A330 (open symbols) during periods of filling in the respective questionnaire (location given in table). dB(B) > dB(A) indicates low-frequency contribution to interior noise level.

the level values in Fig. 6.5 are numerically reduced by 20 dB in order to provide numbers of same magnitude as sound pressure levels, with the objective to construct a balanced input vector, which serves for a neural network analysis (acceleration level of 140 dB equals about 1 g – the reference in Fig. 6.5 is 120 dB). Measurements and input vector are provided by itap GmbH and EADS-CRC [73, 74]. The vibration level is lowest in the front part except for a slight increase in the first galley. The given distribution of vibro-acoustic levels facilitates a certain variability of environmental input to the crew. But of course, the impact of other independent variables has to be taken into account as well.



Abbildung 6.3: dB(A)-levels in cockpit and cabin along aircrafts A330 and A340 from front to aft at respective location during long-haul flights (average in time over flight duration)



Abbildung 6.4: Increase of dB(B) level over dB(A) level along aircrafts A330 and A340 (derived from Fig. 6.3).



Abbildung 6.5: Weighted vibration levels [dB] at same location as in Fig. 6.3. The ordinate scale is shifted by 20 dB to adjust the numbers for training of an artificial neural net (details see text).

6.3.4 Simulator tests

Pre-tests are carried out in the emergency trainer of Austrian Airlines in Vienna (6.6 gives an impression). This simulator has the advantage to provide quite good natural ambience with respect to noise, vibration and motion, but lacks of stable conditions for temperature and humidity. Main simulator tests are carried out in the ACE at Watford, which allows for an excellent stabilization of climatic conditions (except pressure), even at very low humidity [75], but has the disadvantage not to provide motion, which lowers considerably the impression of a natural ambience (impression of the interior of the ACE gives Fig. 6.1).



Abbildung 6.6: "Wired" pilot during pre-test in the cockpit of the simulator of Austrian Airlines

The simulator experiments are carried out by BRE, Medical University Vienna, itap GmbH and Oldenburg University.

Simulator flights could only take place with duration of three to four hours daily due to limitations of the facility. A full $3 \ge 3 \ge 3$ factorial test design is chosen: 3 levels of noise (and vibration, which was derived from the sound signal), 3 levels of temperature, and 3 levels of humidity. The lowest noise level is determined by the background noise of the air conditioning. Because the whole set-up (including test persons) has a considerable relaxation time with respect to the target levels of the climate-parameters (in particular for the humidity), the test design is chosen such that during a simulated 3-hour-flight temperature and humidity are kept constant, whereas noise is adjusted to the given three levels (with a smooth transition which is not noticeable, not even for the supervisors of the experiment).

In short summary, a simulator flight starts with boarding, public welcome address, "start" procedure, one hour "flight" at a given noise level, catering, one hour "flight" at next noise level, catering, one hour "flight" at third noise level, landing procedure. The climatic conditions are kept constant during this session.

Questionnaires are filled in by the passengers and by the crew after each service. The service of the crew includes the distribution of the passengers' questionnaire (see Fig. 6.1).

6.4 Data analysis

The tests results are combined in a data matrix. The rows contain the questionnaire output of each test person in a certain period of activation with corresponding environmental data, medical data and log-data of the respective flight situation. The physical (environmental) data are measured and analyzed according to the technical state-of-the-art, e.g. loudness in Sone, speech interference level SIL, PMV of local thermal comfort, CO and CO₂ with commercial sensors, VOCs with probe tubes and following mass-spectroscopy (according to procedures defined in [76, 77]). The medical (physiological) data of each crew member and test person are analyzed by the Medical University Vienna and summarized in representative (vector) indices [78, 79, 80] characterizing:

- "load" for physiological state and physiological reaction to external stressors
- "imbalance" of subject's physiological status
- alertness, energy, and mood

- subjective health
- performance

Theses psycho-physiological indices are not further discussed in the present paper. Two main approaches are employed in order to relate the environmental condition (including certain intrinsic parameters of each test subject) to the subjective and physiological response of the crew:

- Artificial neural network (ANN)
- Statistical Approach (SA)

The ANN [73, 74] uses an input vector of (partly pre-processed) environmental data and calculates with several intermediate layers an output vector, which is representative for one of the indices mentioned above. The ANN-model works with an error below about 1/2 step-size of the subjective scales (i.e. about $\pm 7\%$ absolute error). Acoustic and vibration data are pre-processed following the known psychoacoustic human response (e.g. text book [81]). The spectral distribution of the cabin interior noise is summarized in third-octave bands and combined with a weighted level of vibration to a so-called multi-spectrum [82], which can in principal take variations in time into account. Since the acoustic signal has more-or-less stationary spectral content in all experiments and is not subject of a defined spectral variation, the time dependency of the multi-spectral representation does not provide additional information to the ANN in the present investigation.

The two ANNs forecast the human response of the investigated clientele quite convincingly, but a sensitivity analysis and a test of general applicability was not yet performed. The ANN approach is not further discussed in the present paper.

The SA uses analysis of variance (ANOVA), principal component analysis (PCA), and correlation analysis to identify relevant environmental and intrinsic parameters which contribute to a certain output of a crew member, e.g. a medical or psychological index or symptom. At a first step it is analyzed, which dependent parameters and data exhibit sufficient variation in answer to the given independent range in the condition of the environment. The analysis produces a reduction of all measured questionnaire data of about 65%. In the next step, a PCA reveals the space of perception of all subjective data, indicates those items which cluster, and gives an estimate, which perceptions are correlated with independent variables, e.g. a certain environmental condition. An ANOVA reveals direct and indirect relations between selected items from the independent and dependant variables. In the following, selected results from the questionnaires with respect to noise and vibration are discussed.

6.4.1 Data from flight tests

A PCA of all data from the questionnaire (i.e. data from all tested crew members, who are at work during flight) displays clear clusters of similar perception, which can be arranged in 11 dimensions related to

- subjective noise-effects (e.g. distraction or annoyance due to noise)
- symptoms (like headache, dizziness, etc.), except those related to dry air or muscle pain
- effects from vibration and motion
- perception of temperature and parameters of local climate (e.g. cold feet)
- motivation, concentration
- perception of air quality
- request to change certain environmental conditions
- communication (incl. intelligibility)
- perceived draft and overall comfort
- symptoms related to dry air
- symptoms related to muscle/ joint pain

Fig. 6.7 gives an example of one plane in the 11-dimensional space of perception with Factor 1 vs. Factor 3, which are obviously related to noise and vibration perception, respectively. Fig. 6.8 demonstrates that noise and vibration do not contribute to the dimensions of the vector space of perception spanned by the Factors 2 and 4. Climate items and numerous symptoms are the dominant factors in this plane.

The correlation of physical environmental parameters like noise level exhibits only a small (but still significant) link to the subjective perception, though the space of perception is well defined and obviously not a product of an arbitrary questionnaire result. A simple linear correlation seems not appropriate to take the changing environmental condition, the mutual interaction, and the complex task load of the crew into account in order to give clear evidence of relations between independent



Abbildung 6.7: Plane of [Factor 1 – Factor 3] of an 11-dimensional space of perception.

and dependant variables. It is therefore necessary to identify significant parameters with the help of ANOVA.

A highly significant dependency (error < 1%) is observed in the subjective development of the air quality (scale from "smelly" to "odorless") in the cabin throughout the whole flight from the beginning (phase 1) to the end (phase 3) for all long-haul flights Vienna-Tokyo-Vienna (Fig. 6.9). The degradation of the perceived air quality is about 15% (i.e. one step of the 7-step scale). The observation does not necessarily indicate an objective change of air quality.

Measurements of independent air quality parameters do not exhibit any degradation. This result points towards an increase of sensitivity of the perception of environmental conditions, which is confirmed by the analysis of a typical symptom (symptom no. 24: "Do you have swollen legs/ feet?") during the three flight phases under the condition of varying noise levels (Fig. 6.10). The perception of this symptom increases significantly with the noise level in the cabin. In particular, the swollen feet become aware at the end of the flight (phase 3) with an increase of 43% under noisy conditions compared to quiet conditions in the beginning (phase 1).



Abbildung 6.8: Plane of [Factor 2 – Factor 4] of an 11-dimensional space of perception.

The result is highly significant with an error < 0.3%.

Another symptom, which has obviously direct impact on the ability to work, is muscle pain in the neck. A pronounced increase of pain with increasing noise level is reported. The symptom is nearly independent from the flight phase, but increases 57% (error < 0.4%) with a noise-level increase of about 8 dB(A). This observation is not yet separated from the work-load in the respective cabin compartment, which are governed by an environmental noise level as depicted in Fig. 6.3.

The discussed results do not take into account the direction of the flight, i.e. if the flight is going from "home" to the remote destination or back "home". Some symptoms like the subjective ability to concentrate (also supported by psychological tests), show a pronounced influence of the flight direction. ANOVA gives a degradation of "concentration" of about 14%, averaged over all flights, when traveling "home" (error < 0.1%).

Insight into the complex interaction of the different parameters is only gained with the help of simulator experiments, which provide investigations under defined con-



Geschätztes Randmittel von c_od

Abbildung 6.9: ANOVA of development of subjective air quality with flight duration in a long-haul flight (beginning of the flight: phase 1; middle of the flight: phase 2; end of the flight: phase 3). The air quality becomes more "smelly" during flight by about 1 step of a 7-step scale (error < 1%).

ditions with a reduced selection of variables.



"at this moment, to what extent are you experiencing swollen or heavy legs/ feet?"(sym_24)

4-step aggregated dB(A) level in cabin

Abbildung 6.10: ANOVA of symptom "swollen feet" (symno.24, scale from 1 to 7) during beginning of flight (phase 1, blue line), in the middle (phase 2, green line) and the end of the long-haul flight (phase 3, grey line), dependant on the noise level dB(A) (aggregated in 4 level-steps). With flight duration and level sym no. 24 becomes aware.

6.4.2 Data from simulator tests

The 3 x 3 x 3 test design allows for the study of variation of two main conditions (Table 6.1): Noise and climate (vibration is derived from the noise signal). With 22 flight crew and 86 cabin crew members serving 544 passengers a PCA exhibits a two-dimensional space of perception, as could be expected from the test design (i.e. variation of the two independent variables noise and climate). One factor relates to all items related to "noise". This result is in agreement with the observation of the flight tests. In contrast to the flight measurement, the dB-level relates well to this subjective factor. The second factor includes all items related to the air quality and subjective views related to temperature conditions.

The test design was such that during a Case I of the simulated flight the noise level increased (step 1, 2, 3) and during Case II the level decreased (step 3, 2, 1) monotonously.

Tabelle 6.1: Test design in ACE simulator with parameter ranges as observed in the cabin. The target values for sound were not realised at each passenger seat, the levels range in about 70-76, 73-78.5, 75-80.5 dB(A)

temperature [°C]	21-22	24 - 25	27-28
relative humidity [%]	5 -10	15 - 20	25-30
sound level step	step 1	step 2	step 3
sound level $[dB(A)]$	70	74	78

ANOVA reveals significant dependencies on perceived noise and vibration and on several symptoms for Case I:

Case I: The noise level has significant impact on (error in brackets)

- level of distraction (3%)
- level of annoyance (1%)
- overall satisfaction (< 1%)
- perception of vibration (< 2%) and movement (< 1%)
- symptoms (< 1%)

lethargy/ tiredness

difficulty in concentration/ remembering swollen or heavy legs/ feet headache

The noise level has no significant impact on any of the previous items in Case II, except for the "perceived volume of noise in the cabin".

An obvious interpretation of this observation is that symptoms in general increase with time, as already identified in the real flight measurements, but consciousness is lowered if the noise level is decreased. If the noise level increases, the crew members become aware of the change of symptoms (Fig. 6.11).



Abbildung 6.11: Scheme to illustrate the trade-off between reported symptoms and change of noise level. Symptoms increase with time, but become aware only if noise level increases as well. Symptoms do not become aware with decreasing noise level

6.5 Conclusion

The impact of environmental conditions like noise and vibration at the workplace of cabin- and flight crew on well-being, performance, and health is identified by measurements in long-haul flights and tests in simulators facilities. Various indices are developed to characterize the human response. A relationship between independent variables (characterizing the environment) and dependant variables (indicators of the human response) is well described by an ANN. Direct inputoutput dependencies of selected parameters reveal a complicated mutual interaction of numerous variables. The noise level exerts significant influence on various symptoms and health indices, in particular when the level increases with time of work. Environmental impact becomes aware with duration of flight.

The authors acknowledge the substantial support of the HEACE project by the European Commission, Proposal No.: GRD 1-2001-40118.

7 Zusammenfassung und Diskussion

Während des Aufenthalts an Bord eines Flugzeuges befinden sich Passagiere und Crew in einer komplexen Umgebung mit multimodaler Anregung. Sie sind dabei Umgebungsbedingungen ausgesetzt, die an Bord der Maschine künstlich erzeugt werden. Nicht zuletzt aufgrund eines prognostizierten Wachstums der Passagierzahlen von jährlich 5% bis zum Jahr 2020 [1] ist neben der Erforschung der unmittelbaren Auswirkungen dieser externen Einflussfaktoren auf die Komfortwahrnehmung auch die gesundheitlichen Beeinträchtigungen durch potentielle Langzeitschäden von aktuellem Forschungsinteresse. Ferner stellt sich die Frage, ob und inwieweit einzelne Parameter an Bord wie Temperatur, Luftfeuchtigkeit, (Unter-) Druck und Akustik untereinander bei der Wahrnehmung interagieren und ob es synergetische Effekte gibt.

Speziell dieser Fragenkomplex wurde für Passagiere im Rahmen des $3\frac{1}{2}$ -jährigen EU Forschungsprojekts" Ideal Cabin Environment" (ICE) mit dem Ziel untersucht, ein neuartiges Vorhersagemodell zu schaffen, welches für physikalische Parameter in der Kabine eine Vorhersage für das subjektiv wahrgenommene Komfortempfinden liefert. Dazu wurden systematisch unter Verwendung von Fragebögen und medizinischen Untersuchungen die Interaktionen der genannten Umgebungsparameter auf Gesundheit und Komfort bei ca. 1.500 Passagieren untersucht, die an insgesamt 36 simulierten Flügen mit je 7-stündiger Dauer teilgenommen haben (Beschreibung in **Kapitel 2**). Wichtiger Bestandteil des Projekts und Aufgabe der Arbeitsgruppe Akustik der Universität Oldenburg war die anschließende Datenanalyse, die zum einen aus einer Datenreduktion und zum anderen in der Ableitung eines geeigneten Modells bestand.

Dazu wurden im Rahmen dieser Arbeit die durch die Projektpartner erfassten und zur Verfügung gestellten Fragebogendaten (Antwortverteilungen) hinsichtlich der geforderten Datenreduktion weiterverarbeitet. Mit Hilfe eines entwickelten und allgemein verwendbaren Verfahrens zur Umskalierung von Antwortskalen (Kapitel 3) wurden die für das Modell relevanten Antworten, die im Verlauf der Vorverarbeitung durch die Partner von einer diskreten und evaluierten Ursprungsskala zunächst auf eine generische normierte Übergangsskala umgewandelt wurden, mittels einer geeigneten Zuordnungsvorschrift gewichtet auf die ursprüngliche Skalenform um- bzw. zurückskaliert (**Kapitel 3.6**). Durch die Abbildung der zu modellierenden Antworten auf die ursprüngliche Fragebogenskala wurde damit die spätere Interpretation der Antwortverteilungen erleichtert.

Parametrisierung von Antwortverteilungen – Kapitel 4

In Kapitel Kapitel 4 folgt die Beschreibung der durchgeführten Datenreduktion mittels Anpassung der Antwortverteilungen an eine statistische Verteilungsfunktion. Um die Antwortverteilungen (jeweils bestehend aus Skalenwert und zugehöriger Häufigkeit) mittels einer standardisierten diskreten Verteilungsfunktion zu parametrisieren, bietet sich insbesondere die Poisson-Verteilung an, welche bereits durch einen einzigen Parameter (λ) vollständig beschrieben ist (Kapitel 4.4.1).

Die Anpassung mittels Minimierung der maximalen Differenz d_{max} zwischen beobachteter und erwarteter kumulierter Häufigkeit (Kapitel 4.4.5) durch geeignete Variation des Parameters λ erwies sich als Bestes der getesteten Verfahren, da das Anpassungskriterium gleichzeitig als Testkriterium im verwendeten Signifikanztest (Kolmogorov-Smirnov D Test oder KS-Test, siehe Kapitel 4.4.4) dient. Für Antwortverteilungen von Fragebogenantworten, die ein Varianz-zu-Mittelwert Verhältnis (VMR) von $R \geq 0,72$ aufweisen, konnte in 95% aller durchgeführten Approximationen eine signifikante Gleichheit zwischen Rohverteilung und angenäherter Verteilung nachgewiesen werden (Kapitel 4.5.3). Abbildung 7.1 zeigt separat für die 29 analysierten Einzelparameter die jeweilige prozentuale Häufigkeit von statistisch signifikanten Anpassungen (KS-Test, β -Fehler < 5%).

Während sich einige Fragenbereiche nahezu vollständig statistisch signifikant approximieren lassen (z. B. Fragen nach wahrgenommener Gesundheit (*perceived symptoms*, siehe Tabelle 4.1), in Abbildung 7.1 rechts mit 'Mean-Sym-xxx' bezeichnet), zeigte sich bei anderen Indizes hingegen eine nur geringe Quote von erfolgreichen Anpassungen mittels der Poisson-Verteilung (siehe 'E-passenger' – fünfter Index v. l. – als Extremfall mit einer Quote signifikanter Anpassungen von 0%).


Abbildung 7.1: Anzahl der Häufigkeiten [in %], für die eine statistisch signifikante Anpassung mittels Poissonverteilung an die gemessene Rohdatenverteilung erreicht wurde (separat für jeden untersuchten Index).
Die Bedeutung der jeweiligen Bezeichnungen der Indizes und Faktoren unterliegt derzeit den Veröffentlichungsrestriktionen des ICE-Projekts.

Neben der Betrachtung des Varianz-zu-Mittelwert Verhältnisses (VMR, Kapitel 4.5.3) kann die Frageformulierung zusätzlich Aufschluss geben, ob eine Antwortverteilung für eine Approximation mittels der Poisson-Verteilung letztlich geeignet ist oder nicht. Der Index 'E-passenger' erweist sich als nicht geeignet: So ist die Antwortverteilung der Beurteilung von (akustischen) Beeinträchtigungen durch andere Passagiere mittels Poisson-Verteilung nicht anpassbar. Es handelt sich um einen Parameter, bei dem sich keine eindeutige Präferenz zu einer Ausprägung einstellt, siehe Abbildung 7.2.



Abbildung 7.2: Gemessene und vorhergesagte Antwortverteilung des Index 'E-passenger' (Beurteilung von (akustischen) Beeinträchtigungen durch andere Passagiere: Skala: 1 = 'keine' bis 4 = 'sehr viele'). Kennwerte der Anpassung: $\lambda_{d_{max}} = 2,06$, statistisch nicht signifikant (KS-Test, β -Fehler $\leq 89\%$), Variance-to-mean ratio: 0, 22.

Im Gegensatz dazu stehen die Fragen zur subjektiv wahrgenommen Gesundheit. Dazu ein Beispiel: Mit dem Index 'Mean-sym-freeze' erhält man eine Antwortverteilung auf die Frage, wie stark die Teilnehmer frieren (Antwortmöglichkeiten 1: *überhaupt nicht* bis 4: *sehr*). Abbildung 7.3 zeigt das Ergebnis der Befragung für Flug 16 bei 21 °C (links) und für Flug 18 bei 25 °C (rechts), jeweils für das 2. Segment (zweite Flugstunde). Die beobachteten Verteilungen wurden mit Hilfe der Poisson-Verteilung parametrisiert. Das Ergebnis ist ebenfalls dargestellt.



Abbildung 7.3: Beobachtete und approximierte Antwortverteilung auf die Frage "Derzeit friere ich ... " mit den Antwortmöglichkeiten 1: überhaupt nicht bis 4: sehr, dargestellt für Flug 16 (21 °C, links) und Flug 18 (25 °C, rechts) (jeweils für das 2. Segment, restliche Parameter: 8000ft. äquivalenter Kabinendruck, 10% rel. Luftfeuchtigkeit, 74dB(A)).

Es zeigt sich, dass in beiden Fällen eine signifikante Anpassung möglich ist (KS-Test, β -Fehler $\leq 5\%$) und zu einem $\lambda_{21 \circ C} = 0,98$ bzw. $\lambda_{25 \circ C} = 0,34$ führt.

Warum Poission-verteilt?

Grundsätzlich scheint die Antwortverteilung einer Gruppe von Probanden auf den ersten Blick nicht Gegenstand eines klassischen Poisson-Prozesses zu sein, wo es z. B. in der Versicherungsmathematik [27] darum geht, die Wahrscheinlichkeiten von seltenen Schadensereignissen (pro Zeitintervall) zu berechnen und vorherzusagen¹. Dennoch ähneln die beobachteten Verteilungsfunktionen denen der Poisson-Verteilung soweit, dass sie sich in großem Maße statistisch signifikant² approximieren lassen. Dazu soll das Ergebnis aus Abbildung 7.3 noch einmal aufgegriffen werden.

Befragt man fiktiv 10 Probanden unabhängig nach ihrem Emfinden des Frierens ("Derzeit friere ich ..." 1: *überhaupt nicht* bis 4: *stark*) bei verschiedenen Temperaturen (z.B. in einem Bereich von 20 °C bis 26 °C), so ist beispielsweise eine innere

¹Aus der berechneten Wahrscheinlichkeit des Schadensereignises werden dann die notwendigen finanziellen Rücklagen der Versicherung zur Deckung bestimmt.

²nachgewiesen über Kolmogorov-Smirnov D Test

7 Zusammenfassung und Diskussion

	Temperatur [°C]						
Teilnehmer	20	21	22	23	24	25	26
1	3	2	2	1	1	1	1
2	3	3	2	2	2	2	1
3	3	2	2	2	1	1	1
4	4	3	3	3	2	1	1
5	2	1	1	1	1	1	1
6	3	2	2	1	1	1	1
7	2	1	1	1	1	1	1
8	3	2	2	2	2	2	1
9	4	4	4	4	3	3	2
10	2	1	1	1	1	1	1
				··· ·			
	beobachtete Häufigkeitsverteilung						
Antwort	20	21			24	25	26
1	0	3	3	5	6	1	9
2	3	4	5	3	3	2	1
3	5	2	1	1	1	1	0
4	2	1	1	1	0	0	0
21 °C		23 °C			25°C		
0		8			8		
0		0				1	
6		6	_		6		
4		4			4		
	2			2			
	0+			0+			
123	4	1	2	34	1	2	34

Repräsentation und damit eine Antwortverteilung gemäß Abbildung 7.4 vorstellbar.

Abbildung 7.4: Fiktives Antwortverhalten von 10 Probanden bei unterschiedlichen-Temperaturen auf die Frage, ob sie frieren (1: *überhaupt nicht* (in grün) bis 4: *stark*, (dunkel blau), oben. Tabellarische Häufigkeitsverteilung (Mitte) sowie für 21 °C, 23 °C und 25 °C grafisch dargestellt (unten).

Unter bestimmten, im Idealfall 'optimalen', Bedingungen (im Beispiel > $26 \,^{\circ}$ C) ist das 'Beschwerdeaufkommen' ("ich friere") eines Teilnehmers bzw. innerhalb der Probandengruppe sehr gering, wohingegen bei zunehmender Verschlechterung (Temperatur sinkt) mehr und mehr Personen das Gefühl des Frierens wahrnehmen und sich dazu entscheiden, eine abweichende Antwort als das Optimum (ich friere '*überhaupt nicht*') zu antworten. In der Analogie zur Versicherungsstatistik

entspricht diese Abschätzung dem "Schadensfall Wolldecke": Nimmt man an, dass Passagiere, die die Frage nach dem Frieren mit '3' oder '4' beantworten, nach einer Wolldecke fragen werden, so ist eine komfortorientierte Fluggesellschaft darauf bedacht, die Passagiere zufrieden zu stellen und eine Wolldecke anzubieten. Wird nun eine Kabinentemperatur von 25 °C eingestellt, so wird in diesem Beispiel (mit 10 Passagieren) bereits eine Wolldecke benötigt. Ab einer Temperatur von ≤ 23 °C werden Wolldecken für zwei und ab 21 °C für drei Passagiere benötigt. Bei Temperaturen ≤ 20 °C steigt der Bedarf auf sieben Wolldecken an.

Die dargestellte Approximation ist definitionsbedingt nur für unipolare Skalen möglich. Für bipolare Skalen existieren bislang ausschließlich erste Überlegungen für ein entsprechendes Anpassungskonzept (mittige Halbierung des Antwortskala mit anschließender separater Anpassung beider Hälften).

Mittels Parametrisierung des vollständigen Antwortverhaltens durch einen einzigen Parameter entsteht eine weitere Möglichkeit der Anwendung, welche in das Vorhersagemodell in ICE eingeflossen ist: Die Interpolation von Verteilungsfunktionen für nicht direkt gemessene Parameter. Dazu nachfolgend zwei Beispiele:

Beispiel 1

Während der Flüge 16 und 18 wurde der Einfluss verschiedener Temperaturen (21 °C und 25 °C) auf den Passagierkomfort untersucht (siehe Kapitel 2.10), u. a. auf das Symptom Frieren ('Mean-sym-freeze', s.o.). Es lassen sich nun die zu 'Mean-sym-freeze' ermittelten λ -Werte über den jeweiligen gemessenen Umweltparameter (Temperatur) aufgetragen. Zusätzlich zu den genannten Werten wird die Beurteilung der 23 °C-Bedingung mit berücksichtigt, welche während der Ruhebedingung zu Beginn der jeweiligen Flüge 16 und 18 eingestellt wurde (Anpassung: $\lambda_{23 \circ C, Flug18} = 0, 62$, gemittelt: $\lambda_{23 \circ C} = 0, 57$). Damit fließt das 'Grundempfinden' der jeweiligen Probandengruppe als Basisbedingung mit in die Vorhersage ein. Je nach Außentemperatur, die am Morgen des Messtages auf die Probanden bei der Anfahrt eingewirkt hat, kann dies leicht variieren.

Das Ergebnis der Regressionsanalyse ist in Abbildung 7.5 links dargestellt.

Zusätzlich kann anhand des Verteilungsparameters λ der vorhergesagte Prozentsatz Unzufriedener (*predicted percentage of dissatisfied*, PPD [35]) berechnet werden (Kapitel 4.6.1). Als 'Unzufrieden' gilt, wer z. B. auf einer 4-stufigen Ratingskala



die Antwort '3' oder '4' gibt, siehe Gleichung 4.7. Die jeweiligen Ergebnisse zeigt Abbildung 7.5 rechts.

Abbildung 7.5: Links: Berechnete λ-Werte für verschiedene Temperaturen (Flug 16 bei 21 °C [*] und Flug 18 bei 25 °C [+]) mit resultierender Regressionsfunktion [- -]. Rechts: Abgeleitete (▲) sowie gemessene Werte (Flug 16 bei 21 °C [+] und Flug 18 bei 25 °C (△)) für den vorhergesagten Prozentsatz Unzufriedener (PPD – Summe der vorhergesagten Wahrscheinlichkeiten für die Antworten '3' und '4' als Ausdruck von Unzufriedenheit). Zusätzlich ist die jeweilige Ruhebedingung beider Flüge bei 23 °C (Segment 1) und die gemeinsame Anpassung mittels Regressionsanalyse (aus den berechneten [- -] und den gemessenen Daten [- ·]) dargestellt.

Führt man eine Regressionsanalyse durch, so erhält man eine Gleichung zur Vorhersage des λ -Werts³ bzw. des vorhergesagten Prozentsatzes Unzufriedener⁴, ausgehend von der eingestellten Temperatur [in °C]. Damit lassen sich nun (nicht gemessene) Zwischenwerte für die Zufriedenheit innerhalb des untersuchten Temperaturbereichs (21 °C bis 25 °C) vorhersagen. Zudem lässt sich – wie erwartet bei signifikant guter Anpassung – eine gute Übereinstimmung zwischen den gemessenen PPD-Werten und den aus der Poisson-Verteilung vorhergesagten PPD-Werten beobachten.

Man erkennt in Abbildung 7.5 für zunehmende Temperaturen ein Absinken der λ -Werte in dem untersuchten Temperaturbereich, während die Werte für den PPD

³Ergebnis der Regressionsanalyse (Polynom 2. Grades) mit R = 1: λ wird berechnet durch $\lambda = 0,023 \cdot T^2 - 1,206 \cdot T + 16,267$ mit R = -0,97, p < 0,05

⁴Regressions analyse (Polynom 2. Grades, R = 1): $PPD_{\lambda} = 0,986 \cdot T^2 - 50,620 \cdot T + 653,684$

ab etwa 24 °C wieder ansteigen. Dieser Verlauf widerspricht jedoch den allgemeinen Erfahrungen, so dass eher damit zu rechnen ist, dass die Anzahl Unzufriedener (in diesem Fall Frierender) mit zunehmender Temperatur weiter abnehmen wird. An dieser Stelle trifft das verwendete Messdesign an seine Grenzen, weil keine weiteren Messugen bei Temperaturen > 25 °C vorgenommen wurden und somit keine weitere Modellierung für höhere Temperaturwerte möglich ist.

Beispiel 2

Als zweites Beispiel ist die Analyse der Beurteilung des Schalls dargestellt. Abbildung 7.6 zeigt (analog zur Temperaturanalyse) das Ergebnis der Approximation des Index 'C-noise', der die Wahrnehmung des Schallpegels durch die Passagiere angibt (siehe Tabelle 4.1). Hierfür wurden die Flüge 22 (bei 64 dB(A)), 7 (69 dB(A)) und 17 (74 dB(A)) analysiert (Beschreibung siehe Kapitel 2.10). Zusätzlich wurde die mittlere Beurteilung der Ruhebedingung zu Beginn des jeweiligen Fluges (55 dB(A)) mit aufgetragen.



Abbildung 7.6: λ-Werte (*) für verschiedene Schallpegel und der sich daraus ableitbare PPD. Zusätzlich ist der gemittelte Wert der Ruhebedingung (Segment 1 mit 55 dB(A), gemittelt über die 3 ausgewerteten Flüge) mit berücksichtigt. Des Weiteren sind die beobachteten Werte für den PPD mit dargestellt (+).

Es zeigt sich, dass die Ruhebedingung von 55 dB(A) von den jeweiligen Probandengruppen ähnlich bewertet wurden (PPD (rechts) schwankt zwischen 13% und 20%). Mit zunehmendem Pegel steigt die Anzahl Unzufriedener auf 50 - 60% bei 74 dB(A). Die berechneten Werte (\blacktriangle) liegen dabei etwas höher als die gemessenen Werte.

Die (lineare) Regressionsanalyse liefert auch hier Vorhersagemöglichkeiten für λ -Werte⁵ und PPD-Vorgersagen⁶. Wie auch bei der Temperatur erlauben die Funktionen nur in dem durch die Tests definierten Bereich zwischen 55 dB(A) und 74 dB(A) eine verlässliche Vorhersage.

Untersuchungen zum Auftreten von Hörschwellenverschiebungen – Kapitel 5

Im Rahmen des ICE-Projekts wurden ebenfalls die Einflüsse der genannten physikalischen Parameter in der Kabine auf die Gesundheit der Passagiere untersucht. Dazu gehörte u. a. die Fragegestellung, ob während eines Langstreckenfluges durch die Exposition mit Kabineninnengeräuschen bei derzeit üblichen Schallpegeln zwischen 70 und 80 dB(A) mit einer zeitweiligen Schädigung des Gehörs (Temporary Threshold Shift - TTS) zu rechnen ist. Es zeigte sich, dass für den Einsatz im Rahmen der Simulatorflüge kein geeignetes Messsystem zur Verfügung stand, was den gestellten Anforderungen an leichter Bedienbarkeit und Schnelligkeit bei gleichzeitig hoher Genauigkeit und zahlenmäßiger Verfügbarkeit entsprach. Deswegen wurde zu Beginn ein Notebook-basiertes, adaptives Messsystems entwickelt und getestet (Kapitel 5.2). Um das Auftreten einer (geringen) Hörschwellenverschiebung von möglicherweise vorkommenden 'normalen' Schwankungen des Hörvermögens im Laufe eines Tages unterscheiden zu können, wurden zunächst die Hörschwellenfluktuationen innerhalb eines Tages durch die Bestimmung der individuellen Ruhehörschwellen bei acht Probanden jeweils morgens und abends an drei verschiedenen Tagen bestimmt (Kapitel 5.3). Bei einem mittleren Expositionspegel von $50 \, \mathrm{dB}(\mathrm{A})$ war von keiner lärminduzierten Hörschwellenverschiebung auszugehen. Es ergab sich eine nicht von Null signifikant⁷ unterschiedliche Verschiebung der Hörschwelle um $0, 3 \, dB \pm 0, 19 \, dB$ (Kapitel 5.3.5).

Im Rahmen einer zweiten Untersuchung wurden vor und zunächst zwei Minuten nach Ende einer 7-stündigen Schalleinwirkung von 80 dB(A) (Darbietung eines Flugzeugkabineninnengeräuschs) die Hörschwellen von 6 Probanden gemessen (Kapitel 5.4). Dabei wurde eine signifikante Hörschwellenverschiebung

 $^{^5 \}mathrm{Ergebnis}$ der Regressionsanalyse mit $R=0,98,\,p<0.05:\,\lambda=0,052\cdot N-2,038$

⁶Regressions analyse: $PPD_{\lambda} = 1,771 \cdot N - 75,949$ mit R = 0,98, p < 0,05

 $^{^7\}mathrm{T}$ Test mit p<0,05

von 3,9 dB \pm 0,51 dB ermittelt (Kapitel 5.4.5). Die in der Literatur dokumentierten Messungen verzeichnen höhere Verschiebungen von ca. 8 dB (Messungen 4 min nach Ende einer 7-stündigen Exposition mit 80 dB-Oktavband-Rauschen um 4 kHz (z. B. [46]). Hypothetische Abschätzungen [54] gehen sogar von einer Verschiebung von bis zu 18 dB aus (2 min (TTS_2 -Wert) nach einer 7-stündigen 80 dB-Exposition bei 4 kHz). Betrachtet man hingegen das Terzpegelspektrum des dargebotenen Signals (realistische Wiedergabe eines real aufgenommenen Airbus A320 Kabinengeräuschs, dargestellt z. B. in Abbildung 5.17), so ist ersichtlich, dass um 4 kHz nur noch ein Schalldruckpegel von ca. 45 dB dargeboten wird. Schalldruckpegel über 80 dB werden nur unterhalb von 200 Hz erreicht, wobei die maximale Hörschwellendifferenz von 6, 2 dB bei 500 Hz ermittelt wurde, gefolgt von 5, 5 dB bei 250 Hz. Diese Werte liegen damit um ca. 2 – 3 dB unterhalb der Literaturwerte, z. B. [46].

Der ebenfalls gemessene Erholungsverlauf (Restitution) mit Messungen 15 bzw. 30 Minuten nach Ende der Exposition mit 1,7 dB \pm 0,50 dB bzw. 0,3 dB \pm 0,48 dB deutet auf einen logarithmischen Zusammenhang der zeitabhängigen Restitution hin, welcher in dieser Form in der Literatur u.a. von Luz und Hodge [38], Kryter [83], Schmidtke [84] oder Praml [85] beschrieben wird. Strasser [55, 86] berichtet ebenfalls von einer Anpassung seiner untersuchten temporären Hörschwellenverschiebungen durch eine logarithmische Funktion, basierend auf einer Berechnung, wie in Abbildung 7.7 dargestellt.

Dabei liefert jede Messung einen definierten Parametersatz a und b in der Anpassungsfunktion $TTS(t) = a + b \cdot \log(t)$. Auf der logarithmischen Zeitachse entspricht der Parameter a einer theoretischen Verschiebung 1 min nach Ende der Exposition (TTS_{1min}) in Dezibel. Der Parameter b definiert das Abklingen der Verschiebung in Dezibel innerhalb einer Dekade auf der Zeitachse. Berechnet man auf Grundlage dieses logarithmischen Zusammenhangs die Parameter a und b für die in dieser Studie ermittelte Restitution aus, so ergibt sich ein Zusammenhang gemäß (7.1).

$$TTS(t) = 4,9-2,97 \cdot \log(t), \ t = [min]$$
(7.1)

mit tals Zeitspanne in Minuten zwischen Ende der Exposition und der Hörschwellenmessung (dargestellt in Abbildung 7.10 als durchgezogene Linie).

Trägt man die von Strasser in [55] dokumentierten Parameterkombinationen a und b für unterschiedliche Messungen gegeneinander auf ([*], Abbildung 7.8) und ergänzt

die Ergebnisse mit dem Wert a = b = 0, was dem theoretischen Wert entspricht, wenn keine Hörschwellenverschiebung (a = 0) und somit auch kein Abklingen vorliegt (b = 0), so lässt sich eine Regressionsanalyse durchführen, die als Ergebnis eine Vorhersagefunktion⁸ für das Abklingen, basierend auf der anfänglichen Hörschwellenverschiebung nach 1 min, liefert (in Abbildung 7.8 mit $[- \cdot]$ dargestellt). Das Ergebnis dieser Untersuchung mit den Parametern a = 4, 9 und b = -2, 97 ist ebenfalls eingezeichnet $[\bullet]$.



Abbildung 7.7: Nach Strasser [86] lassen sich temporäre Hörschwellenverschiebungen mit der logarithmischen Funktion $TTS(t) = a + b \cdot log(t)$ darstellen. Der Parameter *a* entspricht dabei der theoretischen Hörschwellenverschiebung 1 min nach Ende der Exposition (TTS_{1min} , [dB]) und *b* der Erholung (Abklingen der Hörschwellenverschiebung) pro Dekade in Dezibel.

 $^{{}^{8}}b = 0,007a^{2} - 0,645a \text{ mit } R = 0,96, p < 0,05$



Parameter a und b bei unterschiedlichen Restitutionsverläufen mit TTS(t) = a + b * log(t)

Abbildung 7.8: Darstellung von Parameterkombinationen a und b von verschiedenen Hörschwellenverschiebungsuntersuchungen [*], wie sie von Strasser [55] ermittelt wurden. Zusätzlich ist das Ergebnis dieser Untersuchung [●] sowie das Ergebnis der Regressionsanalyse dargestellt ([- ·], unter Berücksichtigung des theoretischen Wertes a = b = 0).

Es zeigt sich, dass die aus der gemessenen Restitution berechnete Parameterkombination in sehr guter Übereinstimmung liegt mit der Erwartungsfunktion, die sich aus den Ergebnissen der Literatur [55] ableiten läßt. Innerhalb des dargestellten Bereichs scheinen die Parameterkombinationen a und b einer Gesetzmäßigkeit zu gehorchen, die durch die angegebene Funktion beschrieben werden kann.

Nach Abschluss dieser Untersuchungen erfolgte der Einsatz des Messsystems zur Detektion einer möglichen temporären Hörschwellenverschiebung durch die Exposition mit Kabineninnengeräuschen zwischen 64 dB(A) und 74 dB(A) im Rahmen der simulierten Langstreckenflüge des ICE-Projekts (Kapitel 5.5). Dazu wurde bei 70 Teilnehmern vor und 30 min nach Ende eines simulierten 7-stündigen Langstreckenflugs mit Schalldruckpegeln von 74 dB(A) die individuelle Ruhehörschwellen gemessen. Es ergab sich eine signifikante Verschiebung⁹ von 0, 72 dB \pm 0, 12 dB und damit eine (leichte) temporäre Hörschwellenverschiebung (TTS), Kapitel 5.5.5.

Abbildung 7.9 zeigt zusammenfassend alle ermittelten Werte der Hörschwellenverschiebungen ($\Delta L_{T,30min}$), die 30 Minuten nach Ende der jeweiligen Exposition (50 dB(A): Messung der Hörschwellenfluktuation, 64, 69 und 74 dB(A)-Kondition: Ergebnisse der Simulationsversuche im ICE-Projekt, 80 dB(A): Untersuchung der Restitution) ermittelt worden sind.



Abbildung 7.9: Ergebnis aller durchgeführten Untersuchungen zur Bestimmung der Hörschwellenverschiebung 30 Minuten nach Ende der jeweiligen Exposition. Dargestellt als Fehlerbalken sind die Standardabweichungen der jeweiligen Einzelmessungen.

Für eine Abschätzung der Höhe der Hörschwellenverschiebung direkt im Anschluss an die Kabinengeräuschexposition $(TTS_2, 2 \text{ Minuten nach Expositionsende})$ dienen die Ergebnisse der vorausgegangenen Diskussionen. Abbildung 7.10 zeigt das Ergebnis der 74 dB(A)-Untersuchung mit 0,72 dB ± 0,12 dB als [\blacklozenge]. Des Weiteren wird das Ergebnis der Untersuchung der Tagesschwankung (mit 50 dB(A) Hintergrundpegel) als durchgehend gepunktete Linie mit der ermittelten Standardabweichung bei

 $^{^{9}\}mathrm{nach}$ T Test, Test auf Gleichheit zu Null, p<0,05

(exemplarisch) 2, 10 und 50 min gezeigt. Als drittes Ergebnis wird der Restitutionsverlauf des 80 dB(A)-Versuchs [■] mit berechneter Regressionsfunktion dargestellt. Die Fehlerbalken geben das 67%-Konfidenzintervall an.



Gemessene Hörschwellenverschiebungen

Abbildung 7.10: Ergebnisse der verschiedenen Hörschwellenverschiebungsexperimente: 50 dB(A) als durchgehend gepunktete Linie mit Standardabweichung exempl. bei 2, 10 und 50 min dargestellt, Ergebnis der 74 dB(A)-Untersuchung [◆] sowie der Restitutionsverlauf des 80 dB(A)-Versuchs [■] mit Regressionsfunktion. Die Fehlerbalken zeigen die jeweilige Standardabweichung des Mittelwerts (1σ-Umgebung) dar. Die Schnittpunkte der Geraden mit der y-Achse sind zur Orientierung mit dem dargebotenen Pegel der jeweiligen Untersuchung markiert.

Es zeigt sich, dass 30 Minuten nach Ende der Exposition mit 74 dB(A), dargestellt als $[\blacklozenge]$, immer noch eine Verschiebung der Ruhehörschwelle von 0,72 dB erkennbar ist, die über der normalen Hörschwellenfluktuation und im Fehlerbereich der Restitutionsmessung bei 80 dB(A) liegt. Nutzt man die in Abbildung 7.8 gezeigte Regressionsfunktion zur Bestimmung der Parameter für die logarithmische Restitutionsfunktion und setzt den gemessenen Wert für die Verschiebung nach 30 Minuten ein, so ergeben sich die Parameter zu a = 6,5 und b = -3,90, was zu einer Hörschwellenverschiebung 2 Minuten nach Ende der Exposition von 5-6 dB führen würde.

Auswirkung physikalischer Umgebungsparameter auf die Flugzeugbesatzung – Kapitel 6

Schwerpunkt des europäischen Forschungsprojekts *Health Effects in Aircraft Cabin Environment* (HEACE)¹⁰ war die Untersuchung der Auswirkung verschiedener Umgebungsparameter auf die Gesundheit und die Arbeitsleistung der Flugbegleiter und Piloten (Crew). Die Ergebnisse der Untersuchung, die zum einen aus bodengebundenen Simulatortests mit systematisch veränderten Bedingungen und zum anderen aus insgesamt 6 realen Langstreckenflugtests (3× Wien - Tokio - Wien und 3× Wien - Neu Delhi - Wien) bestand, wurden zur Entwicklung verschiedener Modelle¹¹ verwendet, welche Vorhersagen für eine optimierte Arbeitsumgebung an Bord eines Flugzeuges ermöglicht.

Die vorliegende Veröffentlichung beschäftigt sich dabei speziell mit den Effekten von Schall und Vibrationen auf die Gesundheit, den Komfort und die Arbeitsleistung der Crew. Eine Untersuchung der Schallpegelverteilung an Bord der verwendeten Airbus-Maschinen A330 und A340 ergab dabei ein Ansteigen des Schallpegels von 10 dB zwischen Cockpit und hinterer Bordküche im Heck (Kapitel 6.3.3). Ebenso wurde durch einen Vergleich von A- und B-bewerteten Schallpegeln gezeigt, dass in der mittleren und hinteren Kabine (ungefähr ab Tragflächeneinsatz) ein hoher Anteil tieffrequenten Schalls auf Passagiere und Crew einwirkt. Eine Faktorenanalyse (Principle Component Analysis (PCA)) aller während der realen Flüge gegebenen Antworten liefert einen 11-dimensionalen Wahrnehmungsraum, wobei Schall und Vibrationen in Faktor 1 und 3 identifiziert werden können (Kapitel 6.4.1). Des Weiteren ist ersichtlich, dass die Luftqualität mit fortschreitender Flugdauer zunehmend als schlechter wahrgenommen wird, obwohl die Messwerte keine Änderungen aufweisen. Dieses deutet auf eine wachsende Sensitivität im Laufe eines Fluges hin. Der gleiche Effekt wird bei subjektiv wahrgenommenen Gesundheitssymptomen beobachtet, besonders unter Einfluss von erhöhten Schallpegeln während des Fluges.

Weiteren relevanten Einfluss auf die Konzentration und die Arbeitsleistung der Flugbegleiter nehmen laut dieser Untersuchung neben der Flugrichtung (Flug weg von

 $^{^{10}}$ www.heace.org, Laufzeit: 2001 – 2005

¹¹Zur Vorhersage wurden zwei unterschiedliche künstliche neuronale Netze (Artificial neural network - ANN) sowie ein Modell, basierend auf der statistischen Datenanalyse, entwickelt.

der Heimatbasis oder zurück) auch das Flugziel ansich und die vorherrschenden tiefrequenten Anteile im Schallpegelspektrum.

Im Rahmen der Auswertung der Simulatorflüge (Kapitel 6.3.4) wurden ebenfalls Schall und Luftqualität als die relevanten Faktoren für die Arbeitsleistung und die Befindlichkeit identifiziert (Kapitel 6.4.2). Basierend auf dem Testdesign der Simulatortests (mit entweder 3-stufig ansteigendem oder abfallendem Schallpegel innerhalb eines Fluges) läßt sich z. B. folgender Zusammenhang zwischen Schallexposition und wahrgenommen Symptomen erkennen: Die Symptome steigen über die Zeit in Anzahl und Stärke an, werden jedoch erst bei zusätzlich ansteigendem Schallpegel bewusst wahrgenommen.

Aufgrund eines im HEACE-Projekt rückblickend nicht optimal gewählten Messdesigns für die Simulatoruntersuchungen wurden weitere, während der realen Langstreckenflüge als relevant identifizierte Parameter nicht untersucht.

Die Untersuchung möglicher Einflüsse dieser offenen Punkte ist ebenfalls nicht mit in das Versuchsdesign des sich zeitlich anschließenden ICE-Projekts aufgenommen worden und bedarf demnach noch immer einer weiteren Auseinandersetzung.

Fazit

Zusammenfassend ergeben sich folgende Ergebnisse:

- Das im Rahmen des Forschungsprojekts ICE entwickelte (allgemeine) Verfahren zur Parametrisierung beobachteter Antwortverteilungen mittels der (diskreten) Poisson-Verteilungsfunktion liefert für den analysierten Datensatz in 95% der betrachteten Antwortverteilungen eine statistisch signifikant gute Anpassung. Damit wird eine erhebliche Datenreduzierung unter Beibehaltung der Detailinformationen zum Antwortverhalten befragter Passagiere ermöglicht. Dieses Verfahren ist universell bei Befragungen mit Antworten auf unipolaren Skalen und einem Varianz-zu-Mittelwert Verhältnis von $R \geq 0.72$ anwendbar.
- 30 Minuten nach Ende einer Exposition mit 74 dB(A) (Kabinengeräusch mit hohem tieffrequenten Anteil) ist eine statistisch signifikante Hörschwellenverschiebung von 0,7 dB ± 0,12 dB messbar, was der Ausprägung einer TTS in geringer Stärke enspricht. Das für die Untersuchung eigens entwickelte adaptive Messsystem zur Bestimmung der individuellen Hörschwelle ist mobil einsetzbar und liefert – auch bei ungeübten Probanden – reproduzierbare Ergebnisse.
- Untersuchungen verschiedener Umgebungseinflüsse auf die Crew eines Flugzeuges innerhalb des Forschungsprojekts HEACE zeigen, dass neben Flugziel und -richtung auch Luftqualität und das Schallpegelspektrum mit einem hohen tieffrequenten Anteil einen Einfluss auf die Arbeitsleistung und das Wohlbefinden aufweisen. Die Symptome steigen über die Zeit in Anzahl und Stärke an, werden jedoch erst bei zusätzlich ansteigendem Schallpegel bewusst wahrgenommen.

Literaturverzeichnis

- [1] Advisory Council for Aeronautics Research in Europe, ACARE Vision 2020, http://www.acare4europe.org/html/documentation.asp (2008).
- [2] H. Mensen, *Entwicklung des Luftverkehrs*, no. 1 in VDI-Buch, Springer Berlin Heidelberg (2007).
- M. A. Bellmann, J. Bastian, R. Weber, Psychophysikalische Untersuchungen zum Gesamtkomfort in Flugzeugkabinen (1/2) - Versuchs- und Rahmenbedingungen -, in Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 31. Deutschen Jahrestagung für Akustik DAGA 2005, München, DEGA e.V., 2005, ISBN: 3-9808659-1-6.
- [4] M. Bagshaw, D. J. Nicolls, Aircraft cabin environment, in Travel Medicine (2nd edition) (2008), pp. 447 – 461.
- [5] E. Parera, Ideal Cabin Environment (ICE) Project: Overview, in Proceedings of the ICE International Aviation Conference, Munich, Germany, ICE consortium (2009).
- [6] CabinAir project, CabinAir consortium, http://projects.bre.co.uk/EnvDiv/cabinair/ (2001 - 2003), EU-gefördertes Projekt innerhalb des GROWTH-Programmes im 5. RTD-Rahmenprogramm.
- [7] FACE consortium, Friendly Aircraft Cabin Environment, http://face.aeronautica.alenia.it (2002 - 2007), EU-gefördertes Projekt innerhalb des GROWTH-Programmes im 5. RTD-Rahmenprogramm.
- [8] HEACE consortium, HEACE-project: Health Effect on Aircraft Cabin Enviroment, http://www.heace.org (2001 – 2005), EU-gefördertes Projekt innerhalb des GROWTH-Programmes im 5. RTD-Rahmenprogramm.
- [9] C. Aizlewood, L. Hamilton, Impact of cabin ventilation rate on aspects of the well-being of aircraft passengers – a laboratory study, in Proceedings of the 26th ICAS Congress, 26th ICAS Congress, Anchorage, Alaska, USA (2008).
- [10] R. Schwab, E. Mayer, DRESSMAN a thermal comfort measurement device for drivers and pilots, in International Journal of Vehicle Design, vol. 42 (2006), pp. 49 – 56.

- [11] G. Grün, A. H. Holm, N. Luks, J. Malone-Lee, M. Trimmel, R. Schreiber, V. Mellert, Impact of cabin pressure on aspects of the well-being of aircraft passengers – a laboratory study., in Proceedings of the 26th ICAS Congress, no. 402, 26th ICAS Congress, Anchorage, Alaska, USA (2008).
- [12] A. Holm, WP3 The test facilities and flight campaign, in ICE Stakeholders Workshop, Holzkirchen, ICE consortium (2007).
- [13] G. Grün, WP2 How does the project address the gaps in knowledge?, in ICE Stakeholders Workshop, Holzkirchen, ICE consortium (2007).
- [14] A. C. Atkinson, A. N. Donev, Optimum Experimental Designs, Oxford University Press (1992).
- [15] NIST/SEMATECH, e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/ (2009).
- [16] M. Trimmel, C. Tischler, K. Trimmel, G. Grün, A. Holm, F. Mayer, I. Baumann, N. Freese, V. Mellert, N. Luks, *Passengers' perception and psychophy*siological response to improved cabin environments in simulated flights (2009), in preparation.
- [17] J. Bortz, N. Döring, Forschungsmethoden und Evaluation f
 ür Human- und Sozialwissenschaftler, Springer, 4. edn. (2006).
- [18] G. Rohwer, U. Pötter, Methoden sozialwissenschaftlicher Datenkonstruktion, Juventa Verlag (2002).
- [19] C. Homburg, Kundenzufriedenheit als strategischer Erfolgsfaktor Einführende Überlegungen, In: Kundenzufriedenheit: Konzepte – Methoden – Erfahrungen, Gabler, Bertriebswirt.-Vlg, 6 edn. (2006).
- [20] L. G. Richards, I. D. Jacobson, Ride quality evaluation 1. questionnaire studies of airline passenger comfort, Ergonomics, 18(2), (1975), 129 – 150.
- [21] I. D. Jacobson, L. G. Richards, Ride quality evaluation 2: Modelling of airline passenger comfort, Ergonomics, 19(1), (1976), 1 – 10.
- [22] L. G. Richards, I. D. Jacobson, Ride quality assessment 3: Questionnaire results of a second flight programme, Ergonomics, 20(5), (1977), 499 – 519.
- [23] I. D. Jacobson, L. G. Richards, Ride quality evaluation 4: Models of subjective reaction to aircraft motion, Ergonomics, 21(7), (1978), 521 – 529.
- [24] E. Mayer, G. Grün, R. Hellwig, A. H. Holm, The new Pressurised Fraunhofer Flight Test Facility offered to the Scientific Cabin Environment Network, in

Proceedings of 1st CEAS European Air and Space Conference 10.-13. September 2007, Berlin, Germany, no. CEAS-2007-468 (2007), pp. 889 – 893.

- [25] G. Grün, A. H. Holm, Passenger well-being in Ideal Cabin Environment, in Proceedings of the IMAPP Conference, IMAPP Conference, Hamburg, Germany (2009).
- [26] I. N. Bronstein, K. Semendjajew, Taschenbuch der Mathematik, B.G. Teubner Verlagsgesellschaft (1991).
- [27] C. L. Trowbridge, Fundamental Concepts of Actuarial Science, Actuarial Education (1989).
- [28] C. Daniel, F. S. Wood, *Fitting Equations to data*, New York: Wiley-Interscience, 5 edn. (1971).
- [29] A. N. Kolmogorow, Grundbegriffe der Wahrscheinlichkeitsrechnung, Springer Verlag (1933).
- [30] F. J. J. Massey, The Kolmogorov-Smirnov test of goodness of fit, Journal of the American Statistical Association, Vol. 46, (1951), p. 70.
- [31] K. Bosch, Elementare Einführung in die angewandte Statistik: Mit Aufgaben und Lösungen, Vieweg+Teubner Verlag, 8 edn. (2005).
- [32] I. M. Chakravarti, R. G. Laha, J. Roy, Handbook of Methods of Applied Statistics, vol. 1, John Wiley & Sons, Inc. (1967).
- [33] W. T. Eadie, D. Drijard, F. E. James, M. Roos, B. Sadoulet, *Statistical Methods in Experimental Physics*, North-Holland, American Elsevier (1971).
- [34] G. Marsaglia, W. W. Tsang, J. Wang, Evaluating Kolmogorov's Distribution, DOAJ-Articles [http://www.doaj.org/oai.article] (Sweden).
- [35] ASHRAE, ANSI/ASHRAE Standard 55-2004: Thermal Environmental Conditions for Human Occupancy (2004).
- [36] H. Dieroff, Lärmschwerhörigkeit, Gustav Fischer Verlag, Jena–Stuttgart, 3. edn. (1994).
- [37] R. Ordonenz, Temporary Changes in Human Hearing Caused by Intense Sounds, Ph.D. thesis, Aalborg University, Aalborg (2005).
- [38] G. Luz, D. Hodge, Recovery from Impulse-Noise Induced TTS in Monkeys and Men: A Descriptive Model, The Journal of the Acoustical Society of America, 49(6), (1971), 1770 – 1777.

- [39] T. Steffens, *Grundlagen Lärm und Lärmschwerhörigkeit*, http://www.uniregensburg.de/Fakultaeten/Medizin/HNO/Laerm.pdf (2007).
- [40] B. Moore, *Psychology of hearing*, Academic Press, San Diego–London–Boston– New York–Sydney–Tokyo–Toronto, 4. edn. (1997).
- [41] M. C. Bottle, S. Charron, H. Bouayad, Temporary threshold and loudness shifts: Frequency patterns and correlations, The Journal of the Acoustical Society of America, 93(3), (1993), 1524 – 1534.
- [42] K. J. Gerhardt, E. L. Hepler, Acoustic-reflex activity and behavioural thresholds following exposure to noise, The Journal of the Acoustical Society of America, 74(1), (1983), 109 – 114.
- [43] D. McFadden, H. S. Plattsmier, Frequency patterns of TTS for different exposure intensities, The Journal of the Acoustical Society of America, 74(4), (1983), 1178 – 1184.
- [44] D. McFadden, Intense sounds may alter the mechanical properties of the cochlear partition, The Journal of the Acoustical Society of America, 72(2), (1983), 447 - 455.
- [45] J. H. Mills, R. W. Gengel, C. S. Watson, J. D. Miller, Temporary changes of the auditory system due to exposure to noise for one or two days, The Journal of the Acoustical Society of America, 48(2), (1970), 524 – 530.
- [46] J. Mills, R. Gilbert, W. Adkins, Temporary threshold shift in humans exposed to octave bands of noise for 16 to 24 hours, The Journal of the Acoustical Society of America, 65(5), (1979), 1238 – 1248.
- [47] J. Mills, W. Adkins, R. Gilbert, Temporary threshold shift produced by wideband noise, The Journal of the Acoustical Society of America, 70(2), (1981), 390 – 396.
- [48] A. Quaranta, P. Portalatini, D. Henderson, Temporary and permanent threshold shift: an overview, Scandinavian Audiology Supplementum, 27(48), (1998), 75 - 86.
- [49] W. Ward, Damage-risk criteria for line spectra, The Journal of the Acoustical Society of America, 34, (1962), 1610 – 1619.
- [50] W. Melnick, Human temporary threshold shift (TTS) and damage risk, The Journal of the Acoustical Society of America, 90(1), (1991), 147 – 154.

- [51] J. Mills, J. Osguthorpe, C. Burdick, J. Patterson, B. Mozo, *Temporary threshold shift produced by exposure to low-frequency noises*, The Journal of the Acoustical Society of America, 73(3), (1983), 918 923.
- [52] P. Kuronen, M. Sorri, R. Pääkkönen, A. Muhli, Temorary threshold shift in military pilots measures using conventional and extended high-frequency audiometry after one flight, International Journal of Audiology, 42(1), (2003), 29 – 33.
- [53] W. Ward, E. Cushing, E. Burns, Effective quiet and moderate TTS: Implications for Noise Exposure Standards, The Journal of the Acoustical Society of America, 59(1), (1976), 160 – 165.
- [54] J. Miller, Effects of Noises on People, The Journal of the Acoustical Society of America, 56(3), (1974), 729 – 764.
- [55] H. Strasser, Traditional Rating of Noise Versus Physiological Costs of Sound Exposures to the Hearing, Volume 66: Biomedical and Health Research, IOS Press, 2nd edn. (2006).
- [56] M. Bagshaw, M. Lower, Hearing loss on the flight deck origin and remedy, The Aeronatical Journal, 106, (2002), 277 – 289.
- [57] T. Küpper, J. Steffgen, P. Jansing, Noise Exposure During Alpine Helicopter Rescue Operations, British Occupational Hygiene Society, 48(5), (2004), 475 – 481.
- [58] 2003/10/EC, Directive 2003/10/EC on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (noise), Official Journal of the European Union (2003).
- [59] G. Békésy, A new audiometer, Acta-Oto-Laryngologica, 34, (1947), 411 422.
- [60] M. Mauermann, G. Long, B. Kollmeier, *Fine Structure of Hearing Threshold and Loudness Perception*, The Journal of the Acoustical Society of America, 116(2), (2004), 1066 1080.
- [61] E. Zwicker, E. Schloth, Interrelation of different oto-acoustic emissions, The Journal of the Acoustical Society of America, 75(4), (1984), 1148 – 1154.
- [62] K. Betke, Hörschwelle und Kurven gleicher Pegellautstärke im ebenen Schallfeld, Ph.D. thesis, Universität Oldenburg – Fachbereich Physik, Oldenburg (1991).

- [63] A. Cohen, J. R. Anticaglia, P. L. Carpenter, Temporary threshold shift in hearing from exposure to different noise spectra at equal dBA level, The Journal of the Acoustical Society of America, 51(2), (1972), 503-507.
- [64] P. A. Hellström, The relationship between sound transfer functions from free field to the eardrum and temporary threshold shifts, The Journal of the Acoustical Society of America, 94(3), (1993), 1301 – 1306.
- [65] R. Hétu, C. Trémolières, Effects of temporal distribution of sound energy on temporary threshold shift produced by intermittent and varying noise exposures, The Journal of the Acoustical Society of America, 61(5), (1977), 1278 – 1287.
- [66] R. Laroche, R. Hétu, A study of the reliability of automatic audiometry by the frequency scanning method (AUDIOSCAN), Audiology, 36(1), (1997), 1–18.
- [67] W. Ward, Temporary threshold shifts and damage-risk criteria for intermittent noise exposures, The Journal of the Acoustical Society of America, 48(2), (1970), 561 – 574.
- [68] T. Yamamoto, K. Takagi, H. Shoji, H. Yoneda, Critical band with respect to temporary threshold shift, The Journal of the Acoustical Society of America, 48(4(2)), (1970), 978 – 987.
- [69] R. Patuzzi, Exponential onset and recovery of temporal threshold shift after loud sound: evidence for long-term inactivation on mechano-electrical transduction channels, Hearing Research, 125, (1998), 17 – 38.
- [70] ISO/Dis/8253.2, Acoustics basic pure tone audiometric test methods, International Organisation for Standardization (1986).
- [71] J. Quehl, Comfort studies on aircraft interior sound and vibration, Ph.D. thesis, Universität Oldenburg (2001).
- [72] DIN EN ISO 7730, Ergonomics of the thermal environment Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria, International Organisation for Standardization (2005).
- [73] M. d'Ischia, A. Brindisi, Evaluation of Psycho-Physiological Response of Civil Aircraft Crewmembers: A Proposal for Modeling, in Proceedings 1st IC-EpsMso, Athens (2005).
- [74] G. Nokas, D. Photeinos, K. Tsokas, D. Tsahalis, A Human Response Model for Aircraft Environment based on Neural Networks, in Proceedings 1st IC-EpsMso, Athens (2005).

- [75] C. J. Skinner, D. J. G. Butler, J. P. Seller, Noise, Vibration and Air Conditioning - Reproduction of Environmental Conditions in an Aircraft Cabin Simulator, in Proceedings 1st IC-EpsMso, Athens (2005).
- [76] DIN EN ISO 16017, Indoor, ambient and workplace air Sampling and analysis of volatile organic components, International Organisation for Standardization (2000).
- [77] VDI 4300 (6) german, Indoor, ambient and workplace air Sampling and analysis of volatile organic components, VDI Guideline (2000).
- [78] M. Trimmel, C. Goger, E. Groll-Knapp, V. Mellert, I. Baumann, N. Freese, R. Weber, M. Bellmann, H. Remmers, M. Grünewald, P. Faulhaber, A. Röder, J. Seller, C. Aizlewood, H. Emms, L. Hamilton, D. Tsahalis, *Modeling the impact* of environmental conditions on comfort, task load and motivation in the mock up and in real flights, in Proceedings 1st IC-EpsMso, Athens (2005).
- [79] M. Trimmel, C. Goger, E. M. Vouk, M. Kritz, A. Klaus, E. Groll-Knapp, *Heart rate of pilots during long haul flights indicate low arousal states*, Psychophysiology, 42.
- [80] M. Trimmel, E. Groll-Knapp, C. Goger, V. Mellert, I. Baumann, N. Freese, R. Weber, M. A. Bellmann, H. Remmers, Impact of environmental conditions on comfort, motivation, task load, physiological activity, health and performance in flight attendants, in Proceedings of the UOEH: Comfort in the Workplace, Kitakyushu (Japan) (2005).
- [81] H. Fastl, E. Zwicker, *Psychoacoustics: Facts and Models*, Springer Verlag (2007).
- [82] M. Bellmann, H. Remmers, A. Röder, Environmental conditions for cockpit and cabin crew members during long-haul flights. From experimental results to a representative vibro-acoustic input vector for ANN and statistical approaches, in Proceedings of 33rd German Annual Conference on Acoustics in Stuttgart (DAGA) (2007).
- [83] K. D. Kryter, The Effects of Noise on Man, Academic Press, New-York-San Francisco-London (1970).
- [84] H. Schmidtke, Handbuch der Ergonomie, Bundesamt f
 ür Wehrtechnik und Beschaffung, Hansa Verlag (1989).
- [85] G. Praml, Umweltmedizin: Lärm, http://www.bgms.de/downloads/004_MFS_A2006-04_Laerm-2HA.pdf (2007).

[86] H. Strasser, H. Irle, Hörschwellenverschiebungen und Restitutionsverlauf nach Impuls- und Dauerschallbelastungen bei der Frequenz maximaler Vertäubung und der unteren und oberen Nachbarfrequenz, in Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 31. Deutschen Jahrestagung für Akustik DAGA 2005, München, DEGA e.V., 2005, ISBN: 3-9808659-1-6.

Stichwortverzeichnis

ACE, 8, 40, 111, 118 ANN, 120 ANOVA, 91, 120 Antwortverteilung, 41

CabinAir, 111

D-dimer, 6 D-optimal design, 15 Datenreduktion, 24, 27 Diskretisierung, 31 Dosimeter, 73 DRESSMAN, 13

FACE, 8, 110 Faktoren, 27 FTF, 9, 40, 61

HEACE, 8, 109

ICE, 5, 61 IDEA PACI, 111 in-flight test, 7, 113 Index, 27, 28, 31, 41, 119 Interleukin-6, 6 Intervallgrenzen, 33 Intervallskala, 28, 31

Kolmogorov-Smirnov-Test, 45

least square fit, 44

Nominalskala, 19 Normierung, 28 Ordinalskala, 19, 28, 31 PCA, 25, 27, 41, 120 PDM, 3, 6, 7, 24, 25, 40 Poisson-Verteilung, 43, 132, 133 PPD, 56, 135 PTS, 62

Quellskala, 33

Ratingskala, 19, 26, 27, 41, 112 recovery curves, 65 reversal points, 69, 104, 105

TTS, 6, 61, 62

VMR - variance to mean ratio, 49 voll faktorielles Design, 15, 119

wobbletone, 68

Zielskala, 33 Zuordnungsvorschrift, 31, 33

Publikationsliste

Übersicht über weitere Veröffentlichungen als Autor oder Koautor – sortiert nach Erscheinungsjahr.

Vorträge – Autor

I. Baumann, M. A. Bellmann, V. Mellert and R. Weber: *Wahrnehmungs- und Unterschiedsschwellen von Vibrationen auf einem Kraftfahrzeugsitz*, Fortschritte der Akustik - DAGA 2001, Hamburg-Harburg, Deutschland, 2001. ISBN: 3-9804568-9-7, 292/293. DEGA e.V., Oldenburg.

<u>I. Baumann</u>, B. Jütz, N. Freese and P. Rong: *Temporäre Hörschwellenverschiebung durch Geräusche in einer Flugzeugkabine?*, Fortschritte der Akustik - DAGA 2007, Stuttgart, Deutschland, 2007, ISBN: 978-3-9808659-3-7, DEGA e.V., Berlin.

<u>I. Baumann</u>, S. Buss, N. Freese, V. Mellert and R. Weber: *Acoustical aspects of travel comfort in the aircraft cabin*, Vortrag auf der Acoustics 2008 in Paris, Frankreich, Abstract in: *The Journal of the Acoustical Society of America*, vol. 123, issue 5, p. 3161 (2008)

<u>I. Baumann</u> and V. Mellert: Distribution of Subjective Assessments in a Controlled Aircraft Environment, in Proceedings of the 2^{nd} International Workshop on Aircraft System Technologies, Hamburg, Deutschland, Shaker Verlag, Aachen 2009.

Vorträge – Koautor

I. Baumann, M. A. Bellmann, S. Buss, N. Freese, E. Groll-Knapp, S. Hallmann, Ch. König, R. Kruse, <u>T. Leitmann</u>, M. Meixner-Pendleton, V. Mellert, H. Remmers, M. Schöls, B. Schulte-Fortkamp, M. Trimmel and R. Weber: *The assessment of load, strain and comfort of flight crew and cabin crew in a cabin simulator - test design,*

Euronoise 2003, Neapel, ISBN: 88-88942-00-9, 2003.

I. Baumann, N. Freese, R. Kruse, V. Mellert and <u>R. Weber</u>: *Psychoacoustic analysis of sound in the cabin of passenger aircrafts*, Euronoise 2003, Neapel, ISBN: 88-88942-00-9, 2003.

I. Baumann, M. A. Bellmann, S. Buss, N. Faulhaber, N. Freese, E. Groll-Knapp, S. Hallmann, Ch. König, R. Kruse, T. Leitmann, V. Mellert, H. Remmers, A. Röder, B. Schulte- Fortkamp, M. Trimmel and <u>R. Weber</u>: *HEACE: Simulation von Schall, Vibrationen und anderer Umgebungsvariablen im Kabinensimulator*, Fortschritte der Akustik - DAGA 2003, Aachen, Deutschland, 2003. ISBN: 3-9808659-0-8, 250/251. DEGA e.V., Oldenburg.

I. Baumann, M. A. Bellmann, S. Buss, N. Freese, <u>E. Groll-Knapp</u>, S. Hallmann, Ch. König, R. Kruse, T. Leitmann, V. Mellert, H. Remmers, B. Schulte- Fortkamp, M. Trimmel and R. Weber: *HEACE: Versuchsdesign zur Erfassung der Belastung, der Beanspruchung und des Komforts von Passagieren und der Crew in einem Kabinensimulator*, Fortschritte der Akustik - DAGA 2003, Aachen, Deutschland, 2003. ISBN: 3-9808659-0-8, 252/253. DEGA e.V., Oldenburg.

I. Baumann, N. Freese, <u>V. Mellert</u> and R. Weber: *Investigation of noise and vibration impact on aircraft crew, studied in an aircraft simulator*, Inter Noise 2004, Prag, Tschechien, 2004. ISBN.

I. Baumann, N. Freese, S. Buss, Ch. König, V. Mellert and <u>R. Weber</u>: *Effects of noise on the comfort of cabin crew studied in an aircraft cabin simulator*, ICA 2004, Kyoto, Japan, 2004.

C. Aizlewood, I. Baumann, M. Bellmann, H. Emms, P. Faulhaber, N. Freese, C. Goger, E. Groll-Knapp, M. Grünewald, L. Hamilton, H. Remmers, A. Röder, J. Seller, <u>M. Trimmel</u>, D. Tsahalis, V. Mellert and R. Weber, *Modeling the impact of environmental conditions on comfort, task load and motivation in the mock up and in real flights*, in *Proceedings 1st IC-EpsMso*, Athens (2005).

I. Baumann, M. A. Bellmann, N. Freese, C. Goger, E. Groll-Knapp, H. Remmers,

<u>M. Trimmel</u>, V. Mellert and R. Weber Impact of environmental conditions on comfort, motivation, task load, physiological activity, health and performance in flight attendants, in Proceedings of the UOEH: Comfort in the Workplace, Kitakyushu (Japan) (2005).

I. Baumann, N. Freese, G. Grün, A. Holm, N. Luks, F. Mayer, C. Tischler, K. Trimmel, <u>M. Trimmel</u> and V. Mellert, *Passengers' perception and psychophysiological* response to improved cabin environments in simulated flights (2009), in preparation.

Eingereichte Veröffentlichung – Koautor

Claire Aizlewood, Ingo Baumann, Andreas Bezold, Johnny Canlas, Jörg Cremers, Michele de Carl, Darius Ereminas, Nils Freese, Gunnar Grün, Andreas Hagen Holm, Lorna Hamilton, Raymond Johnston, Miloslav Jokl, Christos Konstantinidis, Johan Kos, Norbert Luks, Robert Maas, <u>James Malone-Lee</u>, Florian Mayer, Volker Mellert, Vidya Mohamed-Ali, Cinderella Nonoo-Cohen, Earle Perera, Robert Schreiber, Christian Schumacher, Alexander Streit, Catherina Tischler, Karin Trimmel, Michael Trimmel, Jos Vankan, Jaap van Muijden, Giacomo Villi, Jürgen Wenzel, Christian Wolff, Roberto Zecchin and Jane Zuckerman: A study of the health effects of airline passenger cabin environments in simulated 8-hour flights., submitted to The Lancet medical journal, 2009

Danksagung

Die vorliegende Dissertation wurde in der Arbeitsgruppe Akustik der Fakultät V, Institut für Physik, der Carl von Ossietzky Universität Oldenburg angefertigt.

An dieser Stelle möchte ich mich bei all denen bedanken, die mich auf vielfältige Weise bei meiner Arbeit unterstützten, insbesondere:

- bei Herrn Prof. Dr. Volker Mellert und Herrn Dr. Reinhard Weber für die umfangreiche Betreuung und ständige Gesprächsbereitschaft,
- bei Herrn Prof. Dr. Jesko L. Verhey für die Übernahme des Zweitgutachtens,
- bei meinen Kollegen für das angenehme Arbeitsklima in der Akustik sowie für ihre Bereitschaft, jederzeit für kurzfristige Hörtests zur Verfügung zu stehen.
- bei meinen Bürokollegen Herrn Dr. Roland Kruse, der mit seinen umfangreichen Kenntnissen stets zu Diskussionen bereit war, und Frau Daniela Poppinga für die Hilfe bei der statistischen Auswertung sowie der grafischen Aufbereitung der Ergebnisse,
- bei Herrn Ping Rong für die Unterstützung bei der Realisierung der Messsoftware,
- bei Herrn Nils Freese und Frau Berit Jütz für die gute Zusammenarbeit während der ICE-Versuche,
- bei Herrn Dr. Michael Bellmann für die zahlreichen Hinweise und Ratschläge bei der Bearbeitung des Themas,
- bei den beteiligten Partnern in den Projekten HEACE und ICE, insbesondere Herrn Prof. Ing. Dr. phil. habil. Michael Trimmel an der Medizinischen Universität Wien und seinem Team.

Mein Dank gilt nicht zuletzt meinen Eltern, die mich während meiner gesamten Ausbildung uneingeschränkt unterstützten und mir zu jeder Zeit mit Rat und Tat zur Seite standen und mir somit die notwendige Ruhe und Sicherheit für die Erstellung dieser Arbeit gaben.

Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe.

Oldenburg, den 10.08.2009

Lebenslauf

gemäß Paragraph 16, Absatz 3, der Fakultät für Mathematik und Naturwissenschaften der Carl von Ossietzky Universität Oldenburg (11.12.2003)

Name:	Ingo Baumann
geboren am:	23.04.1974
in:	Oldenburg
Staatsangehörigkeit:	Deutsch

Beginn des Studiums an der Universität Oldenburg:WS 1995/96Ende des Studiums:SS 2009

Bildungsgang:

08.1987 - 06.1993	Herbartgymnasium Oldenburg
13.05.1993	Erlangung der Allgemeinen Hochschulreife
07.1993 - 06.1995	Wehrdienst
10.1995 - 08.2001	Studium der Physik an der Universität Oldenburg
08.08.2001	Abschluss Diplom
seit 01.2002	Wissenschaftlicher Mitarbeiter und Doktorand
	in Fak. V, Institut für Physik, Akustik
10.08.2009	Abgabe der Dissertation