

János Heé

Strafe oder Anreiz

**Eine empirische Untersuchung zur
Kundensteuerung in der Airline-Branche**

Strafe oder Anreiz

Eine empirische Untersuchung zur Kundensteuerung in der Airline-Branche

DISSERTATION
der Universität St. Gallen,
Hochschule für Wirtschafts-,
Rechts- und Sozialwissenschaften (HSG)
zur Erlangung der Würde eines
Doktors der Wirtschaftswissenschaften

vorgelegt von

János Heé

von

St. Gallen

Genehmigt auf Antrag der Herren

Prof. Dr. Thomas Bieger

und

Prof. Dr. Torsten Tomczak

Dissertation Nr. 3456

Cuvillier Verlag, Göttingen 2008

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

1. Aufl. - Göttingen : Cuvillier, 2008
Zugl.: St. Gallen, Univ., Diss., 2008

978-3-86727-629-0

© CUVILLIER VERLAG, Göttingen 2008
Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2008
Gedruckt auf säurefreiem Papier

978-3-86727-629-0

Vorwort

Strafe oder Anreiz? Aus betriebswirtschaftlicher Perspektive scheint die Antwort klar zu sein. Kein Unternehmen spricht davon, Kunden zu „bestrafen“, um leistungskonformes Verhalten zu erreichen. In der Airline-Branche haben sich aber Massnahmen zum Branchen-Standard etabliert, die von Kunden als Strafe wahrgenommen werden. Mit der vorliegenden Arbeit werden Einflussfaktoren auf die Verhaltensabsichten eines bestrafte(n) Flugreisenden aufgezeigt. Daraus werden Implikationen für das Dienstleistungs- und Airline-Management sowie für alternative Formen der Kundensteuerung abgeleitet.

Mehrere Personen haben mich während der Dissertationszeit begleitet und unterstützt. Ihnen möchte ich an dieser Stelle herzlich danken.

Mein Doktorvater Prof. Dr. Thomas Bieger verstand es nicht nur, uns Doktoranden beim Skifahren für neue Höhenrekorde zu motivieren, sondern begeisterte mich für das vorliegende Dissertationsthema. Seine kompetente Betreuung legte bereits in einer frühen Phase den Grundstein für diese Dissertation. Ebenso bedanke ich mich bei Prof. Dr. Torsten Tomczak für die freundliche Übernahme des Korreferats. Seine konstruktive Kritik trug wesentlich zum experimentellen Aufbau der Arbeit bei.

Bedanken möchte ich mich bei meinen (Kletter-) Kollegen Andreas Tobler, Felix Lüthi, Matthias Altorfer und Thomas Trawnika, nicht nur für ihre Unterstützung in der *Big Wall*, sondern für diverse Inputs an die Dissertation. Zu grossem Dank bin ich Dr. Torsten von Bartenwerffer und Dr. Torsten Brodt verpflichtet - ohne ihre Korrekturen würde es dieser Arbeit am letzten Schliff mangeln. Die statistischen Auswertungen sind dank Jeannette Oostlanders Korrektur auf methodisch hohem Niveau. Wenn die Software mich bestrafte, waren Michael Kussmauls Programmierkenntnisse unverzichtbar. Auch bedanken möchte ich mich bei Tiziano Pelli als Inspirationsquelle während der gemeinsamen Zeit in der Bibliothek, bei Andri Kofmehl, Bernhard Grämiger, Mathias Öchslin, Rolf Klinkhammer und bei Ruth Flückiger für das Lektorat. Schliesslich möchte ich mich bei meinen Gesprächspartnern aus Praxis und Wissenschaft für ihre Zeit und den wertvollen Input bedanken.

Zoé, Du hast mich mit tausend kleinen Dingen täglich aufs Neue unterstützt und auch mal beide Augen zugeedrückt, wenn es erforderlich war. Dir und meinen Eltern widme ich diese Arbeit. Mit und dank Euch konnte ich diesen spannenden Weg gehen.

Zürich, im Mai 2008

János Heé

Inhaltsverzeichnis

1	Einführung	1
1.1	Ausgangslage und Problemstellung	1
1.2	Stand der Forschung	3
1.3	Forschungslücke	6
1.3.1	Theoretische Lücke	6
1.3.2	Praktische Lücke	7
1.4	Forschungsfrage und Abgrenzung der Arbeit	8
1.4.1	Erkenntnis- und Forschungsobjekt	8
1.4.2	Zielsetzung und forschungsleitende Fragestellung	8
1.4.3	Theoretische und empirische Beobachtungsgrenzen	9
1.5	Forschungsvorgehen und Struktur der Arbeit	10
1.5.1	Wissenschaftstheoretische Ausrichtung	10
1.5.2	Methodisches Vorgehen	12
1.5.3	Forschungsprozess und Aufbau der Arbeit	14
1.6	Zusammenfassung	16
2	Theoretisch-konzeptionelle Grundlagen	17
2.1	Begriffsabgrenzung	17
2.1.1	Dienstleistungsqualität und -marke	18
2.1.2	Massnahmen zur Kundensteuerung	19
2.1.3	Dienstleistungszwischenfall	24
2.1.4	Verhaltensabsichten	26
2.2	Straftheorien im Rechtssystem	30
2.2.1	Notwendigkeit, Sinn und Funktion des Strafrechts	30
2.2.2	Voraussetzungen für Bestrafung gemäss materiellem Strafrecht	31
2.2.3	Funktion der Strafe	31
2.2.4	Privatrecht und Strafe	34
2.2.5	Zusammenfassung	35
2.3	Wahrnehmungs- und Entscheidungsprozesse von Konsumenten	36
2.3.1	Basismodell zur Wahrnehmung	36
2.3.2	Fairness-Beurteilung	40
2.3.3	Attributionstheorie	43
2.3.4	Prospect Theory	44
2.3.5	Zusammenfassung	46
2.4	Strafen im Kontext zwischen Individuen und Organisationen	47
2.4.1	Theorie der Abschreckung und Sättigung	47
2.4.2	Individuen und Organisationen	48

2.4.3	Individuen in der Gruppe	50
2.4.4	Zusammenfassung	51
2.5	Verwandte Konstrukte	51
2.5.1	Erwartungs-Diskonfirmation	51
2.5.2	Kundenzufriedenheit	52
2.5.3	Emotionen bei der Dienstleistungsbegegnung	53
2.6	Provisorisches Wirkungsmodell	54
3	Bestrafung im Airline-Bereich	57
3.1	Grundzüge der Airline-Branche	57
3.1.1	Airline-Geschäftsmodelle	57
3.1.2	Airline-Servicekette	59
3.1.3	Marktsegmentierung nach Flugreisen	62
3.1.4	Kundensteuerung bei einer Flugreise	63
3.1.5	Kundenzufriedenheit und negative Verhaltensabsichten	66
3.1.6	Dienstleistungszwischenfälle auf Flugreisen	67
3.1.7	Wiedergutmachung nach Dienstleistungszwischenfall	68
3.2	Untersuchungsfokus der qualitativen Empirie	69
3.2.1	Untersuchungskontext	69
3.2.2	Erkenntnis- und Forschungsobjekt	70
3.2.3	Forschungsziel	70
3.2.4	Methodenauswahl	70
3.3	E-Mail-Umfrage bei Vielfliegern	71
3.3.1	Erkenntnis- und Forschungsobjekt	71
3.3.2	Forschungsziel	72
3.3.3	Datenerhebung und -auswertung	72
3.3.4	Forschungsergebnisse	72
3.4	Leitfadeninterview mit Experten	73
3.4.1	Erkenntnis- und Forschungsobjekt	73
3.4.2	Forschungsziel	74
3.4.3	Analytisches Raster zur Datenauswertung	74
3.4.4	Datenerhebung	74
3.4.5	Forschungsergebnisse	75
3.5	Beobachtung am Check-In	76
3.5.1	Erkenntnis- und Forschungsobjekt	77
3.5.2	Forschungsziel	77
3.5.3	Analytisches Raster zur Datenauswertung	77
3.5.4	Datenerhebung	78

3.5.5	Forschungsergebnisse	79
3.6	Inhaltsanalyse Airline-Buchungsprozess	80
3.6.1	Erkenntnis- und Forschungsobjekt	80
3.6.2	Forschungsziel	81
3.6.3	Analytisches Raster zur Datenauswertung	81
3.6.4	Datenerhebung	81
3.6.5	Forschungsergebnisse	81
3.7	Empirische Grenzen	82
3.8	Folgerung	83
4	Forschungsmodell und Untersuchungsfokus	85
4.1	Grundzüge der Modell- und Hypothesenbildung	85
4.1.1	Kausalität und Korrelation	85
4.1.2	Moderatorvariablen	86
4.1.3	Experimentelle Hypothesen	86
4.2	Forschungsmodell und Hypothesen	88
4.2.1	Variablenselektion	89
4.2.2	Inhaltlich-theoretische Hypothesen	89
4.3	Untersuchungsfokus der quantitativen Empirie	94
4.3.1	Untersuchungskontext	94
4.3.2	Methodenwahl	95
4.3.3	Erkenntnis- und Forschungsobjekt	96
4.3.4	Forschungsziel	97
4.4	Zusammenfassung	98
5	Empirisch-konzeptionelle Grundlagen	99
5.1	Grundzüge der Experimentalforschung	99
5.1.1	Validitätskriterien für Experimente	99
5.1.2	Web-unterstütztes Experiment	101
5.1.3	Experimentelle Versuchspläne	102
5.1.4	Reaktive Messeffekte	104
5.1.5	Exploration	104
5.1.6	Optimale Stichprobengrösse	105
5.2	Fragebogenentwicklung	108
5.2.1	Einführung	108
5.2.2	Operationalisierung der unabhängigen Variablen	109
5.2.3	Skalenkonfiguration	115
5.2.4	Fragebogenstruktur	119
5.2.5	Pretests	122

5.3	Datenerhebung	123
5.3.1	Flughafen-Umfrage	123
5.3.2	Online-Umfrage	125
5.4	Programme zur statistischen Auswertung	127
5.4.1	Flughafen-Umfrage	127
5.4.2	Online-Umfrage	127
5.5	Zusammenfassung	128
6	Forschungsergebnisse	129
6.1	Deskriptive Datenanalyse	129
6.1.1	Flughafen-Experiment	129
6.1.2	Online-Experiment	131
6.2	Vorgehen und Methoden der statistischen Datenanalyse	132
6.3	Hypothesentest Experiment I	135
6.3.1	Kundensteuerung und Fairness-Beurteilung (h1)	136
6.3.2	Kommunikation und Fairness-Beurteilung (h2)	138
6.3.3	Ticketpreis und Fairness-Beurteilung (h3)	139
6.3.4	Ex ante Image und Fairness-Beurteilung (h4)	141
6.3.5	Fairness-Beurteilung und affektive Reaktion (h5)	143
6.3.6	Emotionale Reaktion und Akzeptanzniveau (h6)	144
6.3.7	Akzeptanzniveau und negative Verhaltensabsichten (h7)	145
6.3.8	Demografische und abhängige Variablen (h_d)	148
6.4	Prüfung Forschungsmodell Experiment I	151
6.4.1	Korrelationen im Kundensteuerungskonzept	151
6.4.2	Interaktions-Hypothesen (h_i)	153
6.4.3	Adaption des Modells für Experiment II	154
6.4.4	Anpassung des Fragebogens für Online-Experiment	155
6.5	Hypothesentest Experiment II	156
6.5.1	Replizierte Hypothesen (h1, h2 und h7)	157
6.5.2	Geschäftsmodell und Fairness-Beurteilung (h8)	158
6.5.3	Akzeptanzniveau und ex post Image (h9)	159
6.5.4	Fairness-Beurteilung und Akzeptanzniveau (h10)	161
6.6	Empirische Grenzen	163
6.6.1	Empirisch bedingte Grenzen	163
6.6.2	Methodisch bedingte Grenzen	164
6.7	Überprüftes Forschungsmodell	166
7	Implikationen für Forschung und Praxis	169
7.1	Implikationen aus dem theoretischen Teil	170

7.1.1	Wahrnehmung der Bestrafungssituation	170
7.1.2	Verhaltensabsichten	172
7.2	Implikationen aus dem empirischen Teil	172
7.2.1	Einflussfaktoren auf Wahrnehmung und Verhaltensabsichten	173
7.2.2	Kundensteuerung in der Airline-Branche	174
7.2.3	Leistungskonformes Verhalten ohne Strafzwang	175
7.3	Implikationen für das Dienstleistungs- und Airline-Management	176
7.3.1	Dienstleistungsmanagement	176
7.3.2	Airline-Management	178
7.4	Kritische Würdigung der Untersuchungsergebnisse	186
7.4.1	Forschungsmodell	186
7.4.2	Methoden-Kombination	186
7.5	Weiterer Forschungsbedarf	187
A	Anhang	189
A.1	Theoretisch-konzeptionelle Grundlagen	189
A.2	Grundzüge der Empirie	190
A.2.1	Beobachtung	190
A.2.2	Experiment	191
A.2.3	Nebenhypothesen	191
A.3	Grundzüge der statistischen Datenanalyse	192
A.3.1	Korrelationen	192
A.3.2	Entscheidungsstatistische Verfahren	194
A.3.3	Varianzanalytische Verfahren in Wirkungsmodellen	197
A.3.4	Effektgrösse	199
A.4	Fragebogen	201
A.4.1	Fragebogen E-Mail-Umfrage	201
A.4.2	Fragebogen Experiment I	202
A.4.3	Fragebogen Experiment II	207
A.5	Forschungsergebnisse	208
A.5.1	Beobachtung am Check-In Flughafen Zürich	208
A.5.2	Inhaltsanalyse des Online-Buchungsprozesses von Airlines	209
A.5.3	Forschungsergebnisse Experiment I	210
A.5.4	Forschungsergebnisse Experiment II	212
A.6	Verzeichnis der Gesprächspartner	219
B	Literaturverzeichnis	221

Zusammenfassung

Knapp kalkulierte Preise und hohe Fixkosten zwingen Airlines, Massnahmen zur Kundensteuerung einzusetzen. Diese Massnahmen können als Push- oder als Pull-Massnahmen gestaltet werden. Je nach Situation während des Serviceprozesses werden die Push- oder Pull-Massnahmen beim Kunden entweder als Anreiz, sich leistungskonform zu verhalten, oder als Strafe wahrgenommen.

Das Forschungsziel der vorliegenden Arbeit ist die Untersuchung des Einflusses von Kundenstrafen auf die Verhaltensabsichten eines Flugreisenden. Hierzu wird basierend auf den theoretisch-konzeptionellen Grundlagen ein provisorisches Wirkungsmodell zur Kundenwahrnehmung und zu den Verhaltensabsichten bei negativ wahrgenommenen Situationen erarbeitet. In einem nächsten Schritt wird das provisorische Wirkungsmodell gestützt auf die theoretischen Grundlagen zum Airline-Bereich, eine Kundenumfrage, Expertengespräche und eine Beobachtung am Check-In zum zentralen A-priori-Wirkungsmodell erweitert. Das zentrale Modell wird experimentell mit einer schriftlichen Umfrage am Flughafen Zürich sowie einer Online-Umfrage überprüft. Das Experiment wird mit der Präsentation verschiedener Szenarien im Serviceprozess durchgeführt. Die Szenarien betreffen das Übergepäck, den Zugang zur Lounge sowie Boarding mit Priorität, ein flexibles Rückreiseticket und ein Upgrade in die nächsthöhere Buchungsklasse. Die Auswertung der quantitativen Empirie erfolgt mit entscheidungsstatistischen und varianzanalytischen Verfahren.

Mit Hypothesentests konnten vier Faktoren identifiziert werden, welche die Kundenwahrnehmung und den Kundenwert beeinflussen: die Massnahme zur Kundensteuerung (Bestrafungsrichtlinie oder Anreiz-Option), die Kommunikation der Kundensteuerung, das Image der Airline und der Ticketpreis. Nehmen Kunden eine Situation als Strafe wahr, werden sie mit negativen Verhaltensabsichten reagieren, insbesondere mit illoyalem Verhalten. Zudem haben Strafmassnahmen, die als unfair beurteilt werden, eine negative Auswirkung auf die Image-Wahrnehmung des Unternehmens.

Strafen werden von Unternehmen als präventive Massnahmen oder als Kompensation für verursachte Kosten resp. entgangene Gewinne gerechtfertigt. Mit Anreiz-Systemen wie bspw. dem Bonusprogramm können Kunden ebenfalls gesteuert werden, ohne negative Verhaltensabsichten auszulösen. Als Management-Implikation der vorliegenden experimentellen Untersuchung folgt die Empfehlung, deutlich kommunizierte Anreiz-Systeme und Bestrafungsrichtlinien kombiniert einzusetzen. Hierfür eignet sich bei Airlines der Internet-Kanal. Bei der Festlegung der Strafgebühr oder des Preises der Anreiz-Option sollten ein faires Verhältnis zum Ticketpreis sowie die Gleichbehandlung aller Passagiere im selben Kundensegment angestrebt werden. Nützlich erweist sich, die *History* eines Kunden zur Kundensteuerung einzusetzen.

Summary

Lean calculated prices and high fixed costs force airlines to implement customer steering measures. These measures can be designed as push or as pull measures. According to the situation during the service process, push or pull measures are perceived by customers as either an incentive to behave service compliant or as a punishment.

The research objective of this paper is the examination of punishment effects on passenger behavioral intentions. Based on the theoretical framework, a provisional research model regarding customer perception and behavioral intentions in negatively perceived situations is being developed. As a next step, this provisional model will be enhanced to the central research model, based on the theoretical fundamentals of the airline sector, a customer survey, expert talks and a scientific observation at the check in. The central research model will be tested experimentally with a written survey at Zurich airport as well as with an online survey. The experiment is being conducted based on the presentation of several scenarios during the service process. These scenarios are related to excess luggage, lounge access as well as boarding with priority, a flexible return ticket and an upgrade to the next higher booking class. The analysis of the quantitative empiricism will be conducted with decision making and variance based statistical procedures.

Testing the hypothesis allowed the identification of four factors, which affect customer perception and customer value: measures to steer customers (punishment policy or incentive option), communication of these measures, airline brand and ticket price. If customers perceive a situation as a punishment, they will react with negative behavioral intentions, particularly with disloyal behavior. Furthermore, punishment measures, which are rated as unfair, have a negative impact on the brand perception of a corporation.

Penalties are being justified by corporations as preventive measures or as balancing in respect of costs caused or lost profits. Steering customers with incentive systems as e.g. the loyalty program does not trigger as many negative behavioral intentions as punishment measures. As a managerial implication of this experimental research it is recommended to deploy both measures, clearly communicated incentive systems and punishment policies, according to the customer involved and the actual situation in the service process. In the airline sector the internet channel qualifies for this purpose. When defining the penalty fee or price of the incentive option, corporations should apply them in a fair relation to the ticket price and the equal treatment of customers of the same segment. In order to steer customers it is beneficial to take the history of a customer into account.

Abbildungsverzeichnis

Abb. 1-1	Forschungsleitende Fragestellung	9
Abb. 1-2	Forschungsprozess und methodisches Vorgehen	15
Abb. 2-1	Kundenverhalten und Kundenzufriedenheit	27
Abb. 2-2	Basismodell zur Wahrnehmung	37
Abb. 2-3	Bestrafung durch instrumentelle Konditionierung	40
Abb. 2-4	Wahrnehmungs- und Entscheidungsprozesse	46
Abb. 2-5	Provisorisches Wirkungsmodell	55
Abb. 3-1	Airline-Servicekette	59
Abb. 3-2	Aufbau und Durchführung der qualitativen Empirie	71
Abb. 3-3	Kundenbefragung zur Fairness-Beurteilung von Richtlinien	73
Abb. 3-4	Dimensionsanalyse zur Kundensteuerung	76
Abb. 3-5	Beitrag der qualitativen zur quantitativen Empirie und zum Forschungsmodell	84
Abb. 4-1	Zentrales A-priori-Forschungsmodell	88
Abb. 4-2	Experimente als Grundlage der quantitativen Empirie	95
Abb. 4-3	Aufbau und Durchführung der quantitativen Empirie	97
Abb. 5-1	Versuchsanordnung im 2x2x2x2 Design (Flughafen-Experiment)	103
Abb. 5-2	Experimentelle Versuchsanordnung	110
Abb. 5-3	Szenarien für Manipulation der Kundensteuerung (UV1)	112
Abb. 5-4	Web-Manipulation mit deutlicher Kommunikation der Anreiz-Optionen	114
Abb. 5-5	Web-Manipulation bei undeutlicher Kommunikation der Richtlinien	115
Abb. 5-6	Struktur Online-Umfrage	120
Abb. 5-7	Durchführung und Auswertung der Experimente	124
Abb. 5-8	Ausfallquote web-gestützter Fragebogen	126
Abb. 6-1	Geflogene Airlines am Tag der Umfrage (Flughafen-Umfrage)	131
Abb. 6-2	Meist geflogene Airlines im 2006 (Online-Umfrage)	132
Abb. 6-3	A-priori-Forschungsmodell Experiment I	135
Abb. 6-4	Korrelationen im Modell in Bezug auf Kundensteuerung	151
Abb. 6-5	Interaktion zwischen Ticketpreis und Kommunikation (Szenario D)	154
Abb. 6-6	A-priori-Forschungsmodell Experiment II	155
Abb. 6-7	Negative Verhaltensabsichten in Szenario B und D (Online-Umfrage)	157
Abb. 6-8	Forschungsmodell nach der Überprüfung	167
Abb. 7-1	Gestaltungsempfehlungen zur Umsetzung der Kundensteuerung	177
Abb. 7-2	Überblick der Effektstärke pro Szenario (Flughafen-Umfrage)	178
Abb. A-1	Primär- und Fehlervarianz als Varianzquellen	191
Abb. A-2	Interaktion zwischen Ticketpreis und Kundensteuerung (Szenario A)	210
Abb. A-3	Interaktion zwischen Image und Kundensteuerung (Szenario B)	211

Abb. A-4	Interaktion zwischen Image und Kommunikation (Szenario A)	211
Abb. A-5	Negative Verhaltensabsichten unter Richtlinie und Anreiz	216
Abb. A-6	Akzeptanz und negative Verhaltensabsichten (Szenario B)	217
Abb. A-7	Akzeptanz und negative Verhaltensabsichten (Szenario D)	218

Tabellenverzeichnis

Tab. 1-1	Stand der Forschung im Bereich Bestrafung	4
Tab. 1-2	Forschungsstand zu relevanten Themen dieser Arbeit	6
Tab. 1-3	Qualitative versus quantitative Forschung	12
Tab. 2-1	Massnahmen zur Kundensteuerung	22
Tab. 2-2	Abgrenzung zwischen kritischen Ereignissen und Bestrafung	25
Tab. 2-3	Negative und positive Verhaltensabsichten	26
Tab. 2-4	Privatrechtliche und öffentliche Strafen im Schweizer Strafrecht	35
Tab. 2-5	Übersicht zu Kognition, Emotion, Motivation und Einstellungen	38
Tab. 2-6	Kognitive Elemente zur Fairness-Beurteilung	42
Tab. 2-7	Attributionen im Kontext von Strafen	44
Tab. 3-1	Marktsegmente Geschäftsreisemarkt	62
Tab. 3-2	Marktsegmente Privatreisemarkt	63
Tab. 3-3	Push- und Pull-Massnahmen bei Airlines	64
Tab. 3-4	Passagierbeschwerden und monetäre Vergütung	67
Tab. 3-5	Kategoriensystem Beobachtung	78
Tab. 3-6	Analytisches Raster für die Inhaltsanalyse	81
Tab. 4-1	Bedeutung und Selektion der Variablen für das Forschungsmodell	89
Tab. 4-2	Hypothesen und involvierte Variablen	93
Tab. 5-1	Optimale Stichprobengrösse für <i>t</i> - und <i>F</i> -Tests (Poweranalyse)	107
Tab. 5-2	Manipulation von Kommunikation, Preis und Image (UV2-UV4)	113
Tab. 5-3	Manipulation des Geschäftsmodells (UV5)	116
Tab. 5-4	Übersicht zur Skalenkonfiguration	117
Tab. 5-5	Kodierung der Skalen	118
Tab. 5-6	Stichprobengrösse (<i>n</i>) pro experimentelle Bedingung (Flughafen-Umfrage)	125
Tab. 5-7	Stichprobengrösse (<i>n</i>) pro experimentelle Bedingung (Online-Umfrage)	126
Tab. 6-1	Demografisches Profil (Flughafen- und Online-Umfrage)	130
Tab. 6-2	Auswirkung von Kundensteuerung auf Fairness-Beurteilung	137
Tab. 6-3	Auswirkung von Kommunikation auf Fairness-Beurteilung	139
Tab. 6-4	Auswirkung vom Ticketpreis auf Fairness-Beurteilung	140
Tab. 6-5	Auswirkung des (ex ante) Image auf Fairness-Beurteilung	142
Tab. 6-6	Zusammenhang zwischen Fairness-Beurteilung und emotionaler Reaktion	144
Tab. 6-7	Zusammenhang zwischen emotionaler Reaktion und Akzeptanz	145
Tab. 6-8	Zusammenhang zwischen Akzeptanz und negativen Verhaltensabsichten	146
Tab. 6-9	Auswirkung der Klassenzugehörigkeit auf Akzeptanz	148
Tab. 6-10	Auswirkung der Klassenzugehörigkeit auf Verhaltensabsichten	149
Tab. 6-11	Auswirkung der Flugerfahrung auf Akzeptanz und Verhaltensabsichten	150

Tab. 6-12	Direkte Auswirkung der Kundensteuerung auf Akzeptanz der Gebühr	152
Tab. 6-13	Korrelation, Signifikanz und Effektstärke in Szenarien A bis D	153
Tab. 6-14	Korrelation, Signifikanz und Effektstärke in Szenarien B und D	156
Tab. 6-15	Auswirkung des Geschäftsmodells auf Fairness-Beurteilung (Online-Umfrage)	159
Tab. 6-16	Auswirkung der Akzeptanz auf (ex post) Image (Online-Umfrage)	160
Tab. 6-17	Zusammenhang zwischen Fairness und Akzeptanz (Online-Umfrage)	162
Tab. 7-1	Framing der Kundensteuerung bei einer Flugreise	183
Tab. A-1	Nebenhypothesen und involvierte Variablen	192
Tab. A-2	Mindestgrößen von signifikanten Korrelationen	194
Tab. A-3	Einteilung der Effektgrößen nach Cohen (1988)	200
Tab. A-4	Qualitative Auswertung Flughafen-Umfrage	212
Tab. A-5	Manipulationscheck für Kundensteuerung (Online-Umfrage)	213
Tab. A-6	Auswirkung der Kundensteuerung auf Fairness-Beurteilung (Online-Umfrage)	214
Tab. A-7	Auswirkung der Kommunikation auf Fairness-Beurteilung (Online-Umfrage)	215

Abkürzungsverzeichnis

A.	Anhang
Abb.	Abbildung
ANOVA	Analysis of Variance
APA	American Psychological Association
AV	abhängige Variable
BA	British Airways
BBC	British Broadcasting Corporation
bspw.	beispielsweise
bzw.	beziehungsweise
ca.	circa
CHF	Schweizer Franken
CRM	Customer Relationship Management
d	Demografie
d.h.	das heisst
DL	Dienstleistung
etc.	und so weiter (et cetera)
et al.	et alii / altera
EUR	Euro
ff.	fortfolgende
FFP	Frequent Flyer Program
h	Hypothese
Hrsg.	Herausgeber
HTML	Hypertext Markup Language
i	Interaktion
i.A.	in Anlehnung
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
IP	Internet Protocol
K., Kap.	Kapitel
LCC	Low Cost Carrier
Neg.	Negative
o.V.	ohne Verfasser
P.b.Korr.	punkt-biseriale Korrelation

Q-Q	Quantil-Quantil
qual. E.	qualitative Empirie
quant. E.	quantitative Empirie
resp.	respektive
S.	Seite
SIA	Singapore Airlines
S-O-R	Stimulus-Organismus-Reaktion
SPSS	Statistical Package for the Social Science
S-R-S	Stimulus-Reaktion-Stimulus
Stand.abw.	Standardabweichung
StGB	Strafgesetzbuch VI
Sum	Summe
Tab.	Tabelle
UA	United Airlines
u.a.	und andere
URL	Uniform Resource Locator
USD	United States Dollar
UV	unabhängige Variable
v.a.	vor allem
vgl.	vergleiche
VIP	Very Important Person
vs.	versus
WOM	Word-of-Mouth (Weiterempfehlungsintention)
z.B.	zum Beispiel
zit.	zitiert

1 Einführung

Dieser einführende Text dient dazu, die Ausgangslage bei Kundenbestrafungen zu skizzieren und daraus die Problemstellung der vorliegenden Arbeit abzuleiten (Kap. 1.1). Unter *Kundenbestrafung* werden (Push-) Massnahmen verstanden, um Kunden zu leistungskonformem Verhalten zu zwingen. Sie werden als Richtlinien in der Einkaufsvereinbarung gestaltet, bei deren Nichtbefolgung eine Strafe erfolgt. Eine alternative Massnahme zur Kundensteuerung sind *Anreiz-Systeme*, die als zusätzliche Produkt-Optionen dem Kunden angeboten werden (Pull-Massnahmen). Im Zentrum dieser Arbeit stehen die Kundenbestrafungen. Der Stand der Forschung zu Bestrafungen im Dienstleistungsbereich wird mit Kapitel 1.2 aufgezeigt. Darauf aufbauend werden sowohl die Forschungslücke als auch die Forschungsfragen und die Abgrenzung der Arbeit abgeleitet (Kap. 1.3; 1.4). Die Beantwortung der Forschungsfragen vor dem Hintergrund der Zielsetzung der Forschungsarbeit erfordert im fünften Schritt die Präzisierung der wissenschaftstheoretischen Ausrichtung und des methodischen Vorgehens (Kap. 1.5). Die Vorstellung der Struktur der Arbeit und die Forschungsplanung sind ebenfalls Bestandteil dieses letzten Kapitels der Einführung.

Einige Konzepte, Theorien und Konstrukte, die in dieser Arbeit verwendet werden, stammen aus dem angelsächsischen Raum. Daher werden stellenweise englische Begriffe eingeführt und kursiv geschrieben, wenn sie nicht allgemein verständlich sind. Für die Zitatform, das Literaturverzeichnis und die (statistischen) Tabellen wird gemäss den Empfehlungen der American Psychological Association (APA) verfahren (vgl. auch Theisen, 2005). Die Zitierweise nach APA eignet sich für Arbeiten in den Wirtschaftswissenschaften (Winkler und McCuen, 1999).

1.1 Ausgangslage und Problemstellung

Unternehmen steuern das Kundenverhalten, indem sie ihnen Richtlinien, z. B. in Form von Gebühren, auferlegen. Damit soll erreicht werden, dass Kunden sich leistungskonform verhalten, da sie die negativen Konsequenzen einer Strafe vermeiden möchten (Prävention). Zusätzlich kann mit monetären Strafen Umsatz generiert werden (Kompensation).

Die Einführung strikter Bestrafungsrichtlinien durch Dienstleistungsunternehmen scheint im Zuge weltweiter Liberalisierung der Märkte und mit Fokus auf Kostenführerschaft ein notwendiges Instrument zu sein, um profitabel zu bleiben. Güter im Dienstleistungsbereich sind vergänglich und verursachen bei Nichtkonsum Kosten. Im Gegensatz zu tangiblen Gütern sind Dienstleistungen nicht lagerfähig, was ein effizientes Management von Angebot und Nachfrage notwendig macht. Eine angedrohte Strafe soll Kunden daher den Anreiz geben, sich unternehmenskonform zu verhalten. Im Falle von monetären Strafen oder Anreiz-Massnahmen als erwerbbar Option können Unternehmen sogar zusätzliche Gewinne erwirtschaften.

Ignorieren Unternehmen Kundenreaktionen auf Bestrafungen, so kann dies in Extremfällen zu lang-

wierigen und imageschädigenden juristischen Auseinandersetzungen führen. Die Firma Blockbuster bspw., ein führender Videoverleih, wurde angeklagt wegen unfairer Gebühren bei zu später Rückgabe von Leihmaterial. Blockbuster hat 17% seines Umsatzes über monetäre Bestrafung realisiert (Horovitz, 2001; Goldstein, 2000). Kritisiert werden auch Unternehmen, die einen erheblichen Teil der Umsätze über Strafgebühren generieren oder tiefe Dienstleistungspreise offerieren und die Reduktion der Margen mit hohen Gebühren wettmachen.

Gemäss Sorell (1994) können Bestrafungen legitimiert werden, wenn das Motiv der Unternehmung Kostenersparnis ist. Die Strafgebühr darf aber nicht zu hoch sein. Kunden sind nicht per Definition im Recht, vor allem dann nicht, wenn sie mit ihrem Verhalten dem Unternehmen schaden. Dienstleistungsunternehmen gehen reaktiv mit dem Thema Bestrafungen um - es gibt wenig Informationen und Hinweise zu Kundenreaktionen und Kundenwahrnehmung bei Bestrafung, zum Management von Strafen, den Effekten der Bestrafung und Alternativen zur Bestrafung. Auch die Forschung beschäftigt sich nur am Rande mit dem Thema Kundenbestrafung (McCarthy und Fram, 2000).

Die Airline-Industrie eignet sich zur Untersuchung von Bestrafungen, weil Richtlinien in vielen Bereichen zum Airline-Industriestandard zählen. Sie werden zwangsläufig eingesetzt, weil jeder Passagier, der sich nicht leistungskonform verhält, hohe Kosten oder fehlende Deckungsbeiträge verursacht (z.B. bei Passagieren, die eine Flugreise nicht antreten). Kunden und Fluggesellschaften stehen während des Serviceprozesses nahe beieinander, emotional wie auch physisch, was die Untersuchung affektiver und kognitiver Reaktionen sowie Verhaltensabsichten aufgrund von Bestrafungen vereinfacht (Verma, 2001).

Bei nicht monetärer Bestrafung ist die Abgrenzung zu kritischen Ereignissen wie schlechtem Service, notwendigen Vertragsbedingungen, Industriestandard oder höherer Macht aus Kundensicht fließend. Die Erarbeitung einer einheitlichen Definition zu Bestrafungsereignissen drängt sich auf. Ein wichtiges Konzept dafür ist das Konzept der negativen kritischen Ereignisse (Dienstleistungszwischenfälle). Auch hier ist die Airline-Industrie als Untersuchungsobjekt geeignet, weil verschiedene Schnittstellen zum Kunden bestehen und somit Unregelmässigkeiten im Serviceprozess unvermeidlich sind (de Coverly et al., 2002). Zudem bestehen unkontrollierbare Variablen wie z.B. das Wetter oder Streiks von Bodenpersonal, welche die Anfälligkeit für Fehler in der Dienstleistungskette weiter erhöhen.

Die Einflussfaktoren auf die Wahrnehmung und auf die Verhaltensabsichten von bestraften Flugreisenden sind in der Forschung bislang nicht untersucht worden. Obwohl Bestrafungsmassnahmen negative Konsequenzen verursachen (negative Verhaltensabsichten von Kunden), werden sie im Dienstleistungsprozess unsystematisch eingesetzt. Alternative Kundensteuerungsmassnahmen zur Bestrafung, wie bspw. Anreiz-Optionen, werden in der Praxis selten eingesetzt. Aufgrund dieser Problemstellung werden in der vorliegenden Arbeit die Folgen und die Effizienz einer als Bestrafung oder als Anreiz wahrgenommenen Kundensteuerung aufgezeigt und Gestaltungshinweise zur optimalen Umsetzung der Kundensteuerung gemacht.

1.2 Stand der Forschung

Die Kundenbestrafung durch Unternehmen und die Reaktion von Kunden auf Bestrafungen ist wenig erforscht (Kim, 2003). Die im Dienstleistungsmarketing durchgeführten Untersuchungen zu Bestrafungen sind relativ jung und betreffen vor allem monetäre Bestrafungen. Beispiele für Strafgebühren verschiedener Dienstleistungsunternehmen sind: Neuausstellung von Flugscheinen oder Übergepäck (Keates, 1998), zu späte Rückgabe von Leihmaterial (o.V., 2003), Bezug an fremden Geldautomaten oder Strafzahlung, weil kein Kauf auf Kreditkarte während einer gewissen Zeitperiode getätigt wurde (Shure, 1996).

Fram hat 1997 eine Untersuchung aus industrieökonomischer Sicht durchgeführt, in der er dreizehn Industrien aufführt, in denen Strafverordnungen benutzt werden (z.B. Autovermietung, Restaurants). Dabei hat Fram mögliche Bestrafungsstrategien analysiert und Szenarien herausgearbeitet, die zu einer erfolgreichen Implementierung dieser Bestrafungsstrategien führen. Fram und McCarthy (1999) haben die industrieökonomische Analyse von Fram (1997) vertieft. Sie haben ermittelt, wie viele Kunden pro Industrie eine Strafgebühr bezahlen mussten und wie hoch die durchschnittliche Strafzahlung in einer Industrie ausfiel. Fram und McCarthy stellten fest, dass Kunden Strafverordnungen prinzipiell akzeptieren. Unternehmen, die sich um ihr Kundenverhältnis kümmern, sollten Richtlinien und Bestrafungsstrategien hinsichtlich der Fairnesswahrnehmung von Kunden untersuchen. Nur auf die Legalität einer Kundenhandlung zu achten, genügt nicht, um nachhaltiges Kundenbeziehungsmanagement zu führen (Berry, 1980). Gemäss Fram und McCarthy sollten Strafen erlassen werden, wenn sie aufgrund von unvermeidbaren Umständen eintreten (z.B. Krankheit oder Todesfall).

McCarthy und Fram (2000) gehen in einer weiteren Untersuchung auf Kundenreaktionen bei Bestrafungen ein. Sie konnten die Wirksamkeit von Strafandrohungen nicht bestätigen: *“The ability of penalties to encourage contract compliance is open to debate”* (McCarthy und Fram, 2000, S. 482). Bestrafungen können aber zu *negativen Verhaltensabsichten* führen wie bspw. tiefere Kundenloyalität oder die Absicht, eine Beschwerde einzureichen. Diese unerwünschten Kundenabsichten können verhindert werden, indem für bestimmte Kundensegmente Strafandrohungen abgeschafft werden. Hierzu muss der Anbieter wissen, welche Kundensegmente mit hoher Wahrscheinlichkeit Verträge erfüllen und nicht bestraft werden müssen. Die Wahrnehmung bezüglich der Bestrafungsfairness von Seiten der Kunden hängt gemäss der Untersuchung von McCarthy und Fram eng zusammen mit der *Kontrollierbarkeit* des Bestrafungsobjektes. Liegt die Bestrafungsursache in einem vom Kunden beeinflussbaren Faktor, so empfindet der Kunde die Bestrafung als fair. Neu eingeführte Strafen, die nicht Industriestandard sind, werden von Kunden als weniger fair empfunden.

Es waren Fram und Callahan, die 2001 erstmals die Fairness-Perspektive und emotionale Reaktionen auf Kundenbestrafungen erforscht haben. Sie untersuchten 44 Kunden mit Bestrafungserfahrung. Im Gegensatz zu den Ergebnissen von Fram und McCarthy (1999) fanden sie in dieser Befragung heraus, dass über die Hälfte der Teilnehmer die erlebten Bestrafungsrichtlinien als unfair

empfangen oder mit der Höhe der Bestrafung unzufrieden waren. Gemäss ihren Ergebnissen sollten Kunden über den Grund der Strafzahlung aufgeklärt werden, um zu verhindern, dass sie der Dienstleistungsunternehmung reine Gewinnabsichten unterstellen. Dem Kunden sollte gemäss Fram und Callahan (2001) erklärt werden, dass mit der Strafzahlung die durch ihn verursachten Kosten gedeckt werden. Strafzahlungen sollten nicht Teil des Geschäftsmodells sein. Es ist ein kompetitiver Vorteil einer Unternehmung, wenn die Richtlinien von Kunden als fair betrachtet werden und umfassend darüber kommuniziert wird. Die Kommunikation umfasst die Erläuterung des Strafgrundes, bei monetären Strafen die Verwendung der Strafzahlung sowie die Art und Weise der Kommunikation. Wiederum die Hälfte aller Befragten versuchte, die Bestrafung mit einer Beschwerde abzuwenden. Nur ein Viertel der Befragten war damit erfolgreich. Hohe Strafkosten verursachten Wut als emotionale Reaktion und resultierten in einer hohen Wahrscheinlichkeit einer Beschwerdeerreichung. Diese erste emotionale und unternehmensschädigende Kundenreaktion kann mit der Aufhebung der bereits ausgesprochenen Strafe aufgehoben oder sogar in eine kundenfreundliche Reaktion verwandelt werden. Kunden, deren Strafe erlassen wurde, reagierten in der Regel wie folgt: *“It was common for respondents to speak highly of the very company that initially penalized them”* (Fram und Callahan, 2001, S. 503). Wenn eine Strafe aber nur teilweise erlassen wurde, fühlte sich der Kunde darin bestätigt, dass er unfair behandelt wurde.

Eine Übersicht zu Forschungsarbeiten im Bereich der Kundenbestrafung ist in Tabelle 1-1 dargestellt.

Tabelle 1-1: Stand der Forschung im Bereich Bestrafung

Autoren	Untersuchungsziel	Hauptaussagen
Fram, 1997	Industrieökonomische Analyse; wirksame Bestrafungsstrategien;	Strafen gewinnen in verschiedenen DL-Industrien an Bedeutung. Strafen müssen der Situation angepasst sein.
Fram und McCarthy, 1999	Industrieökonomische Analyse unter Einbezug Kundenbeziehungsmanagement;	Richtlinien sollten vom Management aktiv gestaltet werden. Ein kritischer Faktor zur Akzeptanz von Kundenbestrafung ist die Ausweichmöglichkeit.
McCarthy und Fram, 2000	Kundenreaktion und Wirksamkeit von Strafen evaluieren;	Kundensegmente reagieren unterschiedlich auf Bestrafung. Unerwünschte Kundenreaktionen können verhindert werden. Wirksamkeit von Strafen ist unklar.
Fram und Callahan, 2001	Fairnesswahrnehmung, emotionale Kundenreaktionen bei Strafe;	Kunden reagieren wütend auf unfaire Strafen. Sie reagieren mit negativer Mund-zu-Mund-Propaganda. Faire Bestrafungsrichtlinien sind kompetitiver Vorteil.
Kim und Smith, 2005	Theoriebasiertes Modell entwickeln zur Wirkung von Strafen auf Kundenzufriedenheit.	Attributionen von Kunden und Straf-Attribute wirken auf die kognitiven Evaluationen und affektive Reaktion. Nach Bestrafung wird die Zufriedenheit besser mit „negative Emotion“ und der „Justice Theory“ erklärt als mit dem Konstrukt „Diskonfirmation“.

Legende: DL = Dienstleistung

Quelle: Eigene Darstellung

Konflikte, die zwischen Unternehmung und Kunden bei Bestrafungen beobachtet wurden, konnten von Fram und Callahan (2001) einem der drei Faktoren zugeordnet werden: Der Kunde war nicht über die Strafe informiert, die Unternehmung verstand die Situation und Wahrnehmung des Kunden nicht oder der Kunde hatte keine Erfahrung mit Strafzahlungen.

Da eine der meistgenannten Kundenreaktion die Kommunikation der erlebten Bestrafung an Freunde und Bekannte ist, sollten Unternehmen ihre Bestrafungsrichtlinien regelmässig überwachen. Bestrafte Kunden erzählen ihren Freunden aber auch, dass ihnen eine Strafe erlassen wurde. Fram und Callahan entwickelten fünf Wege, um Bestrafungsrichtlinien zu verbessern, basierend auf der Erkenntnis, dass Bestrafungsrichtlinien dazu beitragen, Kunden zur Erfüllung von Einkaufsverträgen zu motivieren.

Die erste konfirmativ-quantitative Untersuchung zum Thema *Bestrafung* in Dienstleistungsorganisationen wurde von Kim und Smith (2005) durchgeführt. Im Gegensatz zu den oben erwähnten Arbeiten, die vor allem demografische oder industrieökonomische Charakteristika thematisieren, haben Kim und Smith erstmals ein umfassendes, theoriebasiertes Modell für monetäre Bestrafungen entwickelt und anschliessend empirisch getestet. Sie untersuchten die Effekte von Bestrafungsattributionen sowie die Effekte von Strafattributen (Grad / Schwere, Flexibilität, adäquate Kommunikation) auf die kognitiven Evaluationen und affektive Reaktionen von bestraften Kunden (vgl. Kap. 2.3.3). Forschungsziel von Kim und Smith war es, die Effekte der kognitiven Evaluationen sowie affektive Reaktion von Kunden auf die Beurteilung von monetären Bestrafungen aufzuzeigen.

Resultate der Forschungsarbeit von Kim und Smith zeigen, dass verschiedene Strafattribute in unterschiedlichem Masse die Gerechtigkeitswahrnehmung von Kunden beeinflussen. So hat z.B. die *Schwere* der Bestrafung einen signifikanten und positiven Einfluss auf die Wahrnehmung der Gerechtigkeit, der positiven Diskonfirmation und der negativen Emotionen zur Folge. Daher lohnt es sich für Unternehmen, vor einer Bestrafung jeweils den *Crime Record* eines Kunden zu beachten, um dann die Strafzahlung festzulegen. Flexibilität in der Bestrafungshandhabung hat einen signifikanten Einfluss auf die Gerechtigkeitswahrnehmung. Daher müssen Unternehmen bei der Gestaltung des Bestrafungsmanagements die Grundsätze der *Justice Theory* beachten (vgl. Kap. 2.3.2). Demnach ist nicht nur eine faire Bestrafungshöhe, sondern es sind auch faire Prozesse sowie kommunikativ geschulte Mitarbeiter bei Bestrafungsereignissen nötig.

Bei der Untersuchung der Attributionen zeigen die Resultate von Kim und Smith, dass Kontrollierbarkeit und Stabilität der Strafe einen signifikanten Einfluss auf kognitive Evaluationen und die affektive Reaktion bei Bestraften haben. Daraus schliessen Kim und Smith, dass Unternehmen Bestrafungen klassifizieren sollten, je nach *Kontrollierbarkeit* der Gründe, die zu einer Strafe führten (z.B. zu späte Zustellung einer Kündigung durch Post versus zu spätes Versenden durch Kunden). Kundenprogramme, die die Wahrscheinlichkeit von Bestrafungen verringern (z.B. Versicherungen für verlorene Tickets), sind dabei hilfreich.

Die Gerechtigkeitswahrnehmung, und daher die Fairness einer Strafe, sowie negative Emotionen

sind wichtige Determinanten bei der Bewertung einer Dienstleistung (Kim, 2006). Kim (2006) schlägt vor, dass ein Bestrafungsmanagement eingeführt werden muss, um Kundenwechsel zur Konkurrenz sowie negative Mund-zu-Mund-Propaganda zu verhindern. Zudem sind die Verhaltensabsichten signifikant abhängig vom Lokus der Kausalität (vgl. Kap. 2.3.3). Wenn also Kunden wegen einer Ursache ausserhalb ihrer Kontrolle (z.B. schlechtes Wetter) bestraft werden, werden sie „rebellisch“ reagieren und ihrem Bekanntenkreis davon erzählen.

Abschliessend werden in Tabelle 1-2 die wichtigsten Autoren zu den verwandten und relevanten Themen dieser Arbeit vorgestellt. Erkenntnisse dieser Publikationen werden in den theoretisch-konzeptionellen Grundlagen vertieft (Kap. 2). Die Abgrenzung der verwandten Konstrukte *kritische Ereignisse* und Dienstleistungszwischenfälle in der Airline-Branche wird in Kapitel 3.1 vorgenommen. Theorien zu den Kundenreaktionen und dem Kundenverhalten sind Gegenstand des theoretisch-konzeptionellen Kapitels. Erkenntnisse bezüglich der Dienstleistungsqualität und kritischen Ereignissen im Dienstleistungsbereich sind ebenfalls aufgeführt (Kap. 2.1.1; Kap. 2.5).

Tabelle 1-2: Forschungsstand zu relevanten Themen dieser Arbeit

Wissenschaftlicher Fokus	Autoren
Negative kritische Ereignisse im Airline-Bereich	Stauss et al., 2005; Casado-Diaz und Ruiz, 2002; Verma, 2001; Bejou und Palmer, 1998
Kundenreaktionen auf kritische Ereignisse / Kundenverhalten (Consumer Behavior)	Solomon, 2006; Kroeber-Riel und Weinberg, 2003; Blackwell et al., 2001; Weiner, 2000; Antonides und van Raaij, 1998; Oliver, 1997; Sorell, 1994; Taylor, 1994; Curren und Folkes, 1987
Dienstleistungsqualität / kritische Ereignisse im Dienstleistungsbereich	Bieger, 2003; Fitzsimmons und Fitzsimmons, 2001; Meffert und Bruhn, 2000; Zeithaml und Bitner, 2000; Hays und Hill, 1999; Lovelock, 1992; Normann, 1991; Bitner, 1990; Hart et al., 1990

Quelle: Eigene Darstellung

1.3 Forschungslücke

1.3.1 Theoretische Lücke

Trotz der weiten Verbreitung von Strafrichtlinien, v.a. im Airline-Bereich, gibt es wenige wissenschaftliche Erkenntnisse über ihren Einfluss auf das Kundenverhalten. Die bestehenden Forschungsergebnisse zur Bestrafungsthematik sind nicht geeignet, um die Praxis bei der Implementierung eines ganzheitlichen Bestrafungsmanagements zu unterstützen (Kim, 2006; Fram und Callahan, 2001). Im Airline-Bereich wurden keine wissenschaftlichen Untersuchungen zu Bestrafungen und den Kundenreaktionen sowie dem Kundenverhalten durchgeführt. Insbesondere die emotionale Re-

aktion und negative Verhaltensabsichten nach Bestrafungen eröffnen ein bisher nicht erschlossenes Forschungsfeld.

Für strategische Entscheidungen im Bestrafungsmanagement liegen Forschungsergebnisse vor (vgl. McCarthy und Fram, 2000), aber nur wenige zugängliche Forschungsergebnisse zur Wahrnehmung und emotionalen Reaktion von Kunden, nachdem sie ein Ereignis als Strafe (Richtlinie) oder Anreiz (Option) wahrgenommen haben. Vor allem bei nicht monetären Bestrafungen besteht Forschungsbedarf: *“The examination of other (noneconomic) forms of punitive actions by service organizations and the exploration of motivations behind service firm’s penalty policies [...] do offer interesting avenues for future research”* (Kim und Smith, 2005, S.164). Aus der Kundenperspektive betrachtet, wurde bisher in der Forschung die Rolle der emotionalen Reaktion auf Dienstleistungszwischenfälle vernachlässigt (Schoefer und Ennew, 2005; Bagozzi et al., 1999).

Es überrascht nicht, dass bisher kein Modell zur Taxonomie von Bestrafungen und die Gegenüberstellung zu Anreizsystemen als Kundensteuerungsmassnahmen entwickelt wurde. Strafen könnten bspw. in zwei Gruppen eingeteilt werden: Strafen, die einen Nutzen versprechen (späte Rückgabe eines Leihgegenstandes), und Strafen, die keinen Nutzen bringen (*Downgrading* im *Frequent Flyer Program*). Ferner lassen sich Situationen evaluieren, bei denen die Effizienz von Anreizmechanismen gegenüber Richtlinien als Kundensteuerungsmassnahmen höher ist.

Demnach werden mit dieser Arbeit (1) die Lücke der fehlenden Definition zu Bestrafungen (im Airline-Bereich) und (2) die Wahrnehmung von Strafen und Anreiz-Massnahmen sowie die Verhaltensabsichten und emotionalen Reaktionen bei Kunden untersucht.

1.3.2 Praktische Lücke

Auch die Praxis setzt sich nicht mit der Thematik auseinander: *“Our inability to uncover substantial prior research in this area, along with the unwillingness of executives to discuss this topic indicates that executives are not giving their penalty policies the attention that they warrant”* (McCarthy und Fram, 2000, S. 495, und ähnlich: Kim, 2006). Dienstleistungsunternehmen benützen heute vor allem verschiedene Arten von Richtlinien mit Bestrafungen, ohne sich vollständig über die - meist negativen - Effekte auf die Kundenwahrnehmung dieser Strafen bewusst zu sein. Gemäss Erkenntnissen der qualitativen Empirie besteht bei Airlines keine einheitliche Strategie, wie mit Anreiz-Massnahmen / Optionen und Bestrafungen während der verschiedenen Stufen des Serviceprozesses umgegangen werden soll (vgl. Kap. 3.4). Viele Dienstleistungsunternehmen setzen beide Instrumente gleichzeitig ein, ohne sich über die Wirkungen und die Effizienz im Klaren zu sein. Dementsprechend besteht auch keine Handlungsanweisung, um den verschiedenen negativen Verhaltensabsichten von Kunden adäquat zu entgegenen.

Die praktische Lücke bezüglich den Kundensteuerungsmassnahmen wird mit der qualitativen Forschung zum Konstrukt *Bestrafung* und *Anreiz-Massnahmen* (3) sowie den Implikationen zur praktischen Umsetzung effizienter Kundensteuerung geschlossen (4).

1.4 Forschungsfrage und Abgrenzung der Arbeit

1.4.1 Erkenntnis- und Forschungsobjekt

Das Forschungsobjekt dieser Arbeit sind „Sanktionen“ (*Kundensteuerung mit Richtlinien*), die von Kunden als Bestrafung wahrgenommen werden, und die Kundensteuerung mit „Anreiz-Optionen“. Sie werden am Beispiel von Airlines als Dienstleistungsunternehmen untersucht. Weil der Transport von Flugpassagieren ein *high contact* Service ist (vgl. Kap. 3.1), können verschiedene Situationen und Kundensteuerungsmassnahmen gegenübergestellt werden (vgl. Kap. 2.1.2). Als Massnahme zur Kundensteuerung stehen Richtlinien den Anreizsystemen mit „Belohnungen“ gegenüber. Um eine ganzheitliche Sicht auf Kundensteuerungsmassnahmen zu gewährleisten, werden Anreiz-Massnahmen mit Optionen als Alternative zu Strafen empirisch untersucht. Die Begriffe *Bestrafung* und *Anreiz-Option* werden somit als definitorische „Spiegelbilder“ betrachtet. Sie werden unter Kundensteuerungsmassnahmen subsumiert.

Um Einblick in das Konstrukt *Bestrafungen* zu erhalten, wird vorerst Rückgriff auf Theorien der Rechtswissenschaften genommen: wozu Strafen dienen und was ihr Hintergrund ist (vgl. Kap. 2.2). Um das Forschungsmodell aufzubauen, werden die Ergebnisse des Dienstleistungsmarketings zu kritischen Ereignissen und Kundensteuerungsmassnahmen eingesetzt (vgl. Kap. 2.1). Dabei erfordert der Fokus auf die Kundenwahrnehmung und -einstellung sowie die emotionalen Reaktionen resp. Verhaltensabsichten den Einbezug ökonomischer, sozialpsychologischer und soziologischer Forschungserkenntnisse (vgl. Kap. 2.3; 2.4).

1.4.2 Zielsetzung und forschungsleitende Fragestellung

Ziel dieser Arbeit ist es, Bestrafungen aus Sicht von Flugreisenden zu evaluieren und darauf aufbauend ein verhaltenswissenschaftliches Erklärungsmodell zur Wirkung von Strafen und Anreizen auf die Reaktion von Kunden aufzubauen. Es wird ein Beitrag zur Theorie geleistet: Einflussfaktoren auf die affektive und kognitive Reaktion von Kunden nach einem Bestrafungsereignis sowie der Zusammenhang zwischen dieser Reaktion und den Verhaltensabsichten. Mit den Implikationen für das Dienstleistungsmanagement werden effiziente Steuerungsmassnahmen, um leistungskonformes Verhalten zu erreichen, in verschiedenen Situationen aufgezeigt. Dabei steht die Beziehung zwischen Dienstleistungsunternehmen und Kunden in verschiedenen Situationen im Vordergrund.

Die wissenschaftliche Zielsetzung der Arbeit gliedert sich in die konzeptionelle Aufbereitung und die theoretische Fundierung von Bestrafungen sowie die empirische Überprüfung des abgeleiteten Modells zur leistungskonformen Kundensteuerung (vgl. Kap. 2.6). Der empirische Beitrag liegt in der Überprüfung der in Kapitel 4 an der Realität abgeleiteten Hypothesen (Kromrey, 2006). Das Wissenschaftsziel für die Praxis besteht in der managementpraktischen Anwendung der theoretischen Aussagen für die Kundensteuerung bei Airlines und Dienstleistungsunternehmen. Aus den Zielsetzungen ergeben sich die forschungsleitende Hauptfrage und die Unterfragen a.1 bis c.2, die in Abbildung 1-1 dargestellt sind.

Abbildung 1-1: Forschungsleitende Fragestellung

Forschungsziel	Welchen Einfluss haben Kundenbestrafungen auf die Verhaltensabsichten eines Flugreisenden?	
Theoretischer Bezugsrahmen	a.1	Wie nehmen Individuen Bestrafungssituationen wahr und worauf basiert die Verhaltensreaktion?
	a.2	Was sind die theoretischen Beiträge zur Bestrafungsthematik aus Sicht der Jurisprudenz, Psychologie, Ökonomie und Soziologie?
Empirie	b.1	Welche Faktoren beeinflussen die Wahrnehmung und die Verhaltensabsichten verschiedener Kundensegmente? (qual. und quant. E.)
	b.2	Wie und auf welcher Service-Stufe steuern Airlines ihre Passagiere? (qual. E.)
	b.3	In welchen Situationen kann leistungskonformes Passagierverhalten ohne Strafzwang, mit Anreiz-Optionen, erreicht werden? (quant. E.)
Implikationen	c.1	Beitrag an die Wissenschaft: Wirkungsmodell zur Kundensteuerung.
	c.2	Implikationen für das Dienstleistungs- und Airline-Management zur optimalen Steuerung des Kundenverhaltens in verschiedenen Situationen.

Legende: qual. = qualitative, quant. = quantitative, E. = Empirie;

Quelle: Eigene Darstellung

Die theoretisch zu beantwortenden Forschungsfragen (a.1, a.2) werden in Kapitel 7.1, die empirisch zu beantwortenden Forschungsfragen (b.1 - b.3) in Kapitel 7.2 beantwortet. Die Implikationen werden bezüglich der Forschungsfrage c.1 in Kapitel 6.7 und für c.2 in Kapitel 7.3 gemacht. Die Forschungsfragen dienen der Schliessung der in Kapitel 1.3 aufgezeigten Forschungslücke, indem:

1. Bestrafungen aus verschiedenen wissenschaftlichen Disziplinen betrachtet (a.1, a.2) und erstmalig Bestrafungen im Airline-Bereich erforscht werden (b.2);
2. Nicht nur die emotionale Reaktion und kognitive Evaluation von Kunden, sondern die Verhaltensabsichten untersucht werden (a.1, b.1);
3. Auf Basis qualitativer Forschung Informationen über das Konstrukt *Bestrafung* und *Anreiz-Massnahmen* im Airline-Bereich gewonnen werden (b.3);
4. Auf Basis der quantitativen empirischen Untersuchung praktische Implikationen zur effizienten, leistungskonformen Kundensteuerung abgeleitet werden (c.1, c.2).

1.4.3 Theoretische und empirische Beobachtungsgrenzen

Die Vielfalt der formulierten Zielsetzungen und die zu berücksichtigenden Aspekte und Themengebiete erfordern eine Abgrenzung des Analysegegenstandes.

In dieser Arbeit wird von Situationen ausgegangen, bei denen die Unternehmung versucht, Kunden im Serviceprozess zu steuern und dabei Bestrafung oder Anreiz-Massnahmen einsetzt. Das zu er-

arbeitende Wirkungsmodell beschränkt sich dabei auf die Abbildung von Verhaltensabsichten von Kunden mit negativen emotionalen Reaktionen und entsprechenden kognitiven Prozessen bei verschiedenen Kundensteuerungsmassnahmen. Im Zentrum des theoretischen Teils stehen also weder die Markenwahrnehmung, Fehler im Dienstleistungsprozess noch das konkrete Kundenverhalten nach einem Ereignis.

Die empirische Forschungsarbeit ist aus forschungsökonomischen Gründen eingegrenzt auf Flugreisende vom Flughafen Zürich (Lokal- und Transitpassagiere). Da sich die Erwartungshaltungen an eine Flugreise zwischen verschiedenen Kundensegmenten und Kulturen unterscheiden, wird bei der Datenerhebung darauf Rücksicht genommen und eine international ausgerichtete, web-unterstützte Umfrage als Ergänzung durchgeführt. Im Gegensatz zu emotionalen Reaktionen sind kognitive Evaluationen schwer messbar. Um die im Zusammenhang mit Bestrafungsereignissen relevanten kognitiven Prozesse abzubilden, wird Rückgriff auf Erkenntnisse der empirischen Ökonomie, Soziologie und Psychologie genommen (vgl. Kap. 2.3; 2.4).

1.5 Forschungsvorgehen und Struktur der Arbeit

1.5.1 Wissenschaftstheoretische Ausrichtung

Die vorliegende Arbeit geht von einem Praxisproblem aus, für das aus praktischer als auch theoretischer Sicht unzureichende Erklärungen vorliegen. Die Forschungsfragen sind von Beginn an auf das Erkennen eines konkreten Praxisproblems ausgerichtet (Ulrich, 1981). Aus inhaltlich-konzeptioneller Perspektive wird das Dienstleistungsmarketing als Forschungsdisziplin der Betriebswirtschaftslehre und als angewandte Wissenschaft aufgefasst (Ulrich, 2001). Sie hat die Aufgabe, konzeptionelle Entscheidungshilfen für die Praxis zu entwickeln. Wegen des engen Bezugs zum Konsumentenverhalten wird auch auf die Käuferverhaltens- oder Konsumentenforschung Rückgriff genommen. Diese zielt darauf ab, das Verhalten von Kunden zu verstehen und zu erklären sowie Handlungsempfehlungen zur Beeinflussung des Käuferverhaltens abzuleiten (Kuss und Tomczak, 2004; Kroeber-Riel und Weinberg, 2003; Trommsdorff, 2003; Gröppel-Klein, 2001).

Die amerikanisch geprägte Marketingwissenschaft ist überwiegend vom *positivistischen Forschungsparadigma* dominiert (Göthlich, 2003; Hunt, 2003; Easton, 2002). In diesem Paradigma wird Erkenntnis aus objektiv feststellbaren und verifizierbaren Tatsachen gewonnen. Es unterstellt die Existenz natürlicher Gesetzmässigkeiten für die Sozialwissenschaften und übernimmt das Ideal und die Methodik aus der Physik, wonach nur eine wahre Natur der Realität existiert (Ulrich, 2001). An diesem positivistischen Paradigma gibt es prominente Kritik (z.B. durch Feyerabend, 1975)¹, aber kein einheitliches Gegenprogramm, sondern eine Fülle an Konzepten, die unter *Konstruktivismus* oder *Relativismus* subsumiert werden. Auf der philosophischen und methodologischen Ebene

¹ Für Feyerabend (1975) simplifiziert die wissenschaftliche Ausbildung des kritischen Rationalismus durch eine Kette von Festlegungen, Ausgrenzungen und Definitionen. Sie lässt den Eindruck einer feststehenden Welt und einer feststehenden Wissenschaft erzeugen.

stehen demnach positivistische-empirische und relativistische-konstruktivistische Ansichten gegenüber (Peter und Olson, 1983).

Im Konstruktivismus wird postuliert, dass die Realität, mit der sich die Wissenschaft beschäftigt, nicht objektiv gegeben, sondern kognitiv erzeugt ist. Wissenschaftliche Erkenntnis ist beim Konstruktivismus demnach möglich. Aber sie ist keine Erkenntnis der Wirklichkeit, sondern ein Teil der Konstruktion der Wirklichkeit. Wissenschaftliche Erkenntnis ist stets subjektiv, zeitgebunden und vom spezifischen Kontext der Untersuchung abhängig (Jensen, 1999; Schmidt, 1991; von Glasersfeld, 1985; Deshpandé, 1983). Der Relativismus besagt, vereinfacht ausgedrückt, dass etwas relativ zu etwas anderem ist. Die Begründung von Aussagen bzw. Handlungen ist somit immer nur unter der Voraussetzung von Prinzipien möglich, die selbst keine universelle Gültigkeit haben (Carrier, 1995b; Anderson, 1983; Peter und Olson, 1983). Anhänger des Relativismus zeigen im Gegensatz zu *Positivisten* mehr Toleranz gegenüber alternativen Forschungsmethoden. Folgende Auffassungen werden somit im relativistischen-konstruktivistischen Wissenschaftsverständnis vereint: Die Realität ist relativ und vielfältig, Theorien haben keine universelle Gültigkeit, die wissenschaftliche Erkenntnis ist stets subjektiv geprägt und die Wissenschaft ist als sozialer Prozess zu verstehen und durch Interaktion geprägt (Jensen, 1999; Carrier, 1995a; Deshpandé, 1983).

Durch mangelnde theoretische Durchdringung - im Falle des vorliegenden Forschungsobjekts - führt der kritische Rationalismus nur zu einem beschränkten Erkenntnisgewinn (Tomczak, 1992). Lepplin (1986) und Hunt (1983) leiten aus der Kritik am kritischen Rationalismus die wissenschaftstheoretische Konzeption des *wissenschaftlichen Realismus* ab. Es ist eine Annäherung an die relativistisch-konstruktivistische Position, wobei die Objektivitätsprämisse eingehalten wird. Dieses Konzept beruht auf der Annahme, dass die Hypothesenprüfung an der Realität positiv ausfallen kann. Damit lassen sich zwar keine universellen Wahrheiten folgern, aber dafür tatsächlich bestehende Kausalzusammenhänge. Zudem sind Messfehler zuzulassen und eine induktive Schlussweise möglich (Witte, 1977). Der wissenschaftliche Realismus versteht Theorien als Zusammenhang zwischen latenten Konstrukten, die nur mit fehlerhaften Indikatoren gemessen werden können (Homburg, 1989; Anderson und Gerbing, 1982). Für das vorliegende Forschungsvorhaben könnte die Ausrichtung am wissenschaftlichen Realismus insofern beitragen, weil es sich beim Sachverhalt erstens um hypothetische Konstrukte handelt, welche sich der direkten Messung entziehen, und zweitens, weil die Kontextfaktoren nicht vollständig kontrollierbar sind.

Trotzdem wird in dieser Arbeit wissenschaftstheoretisch vom relativistisch-konstruktivistischem Forschungsverständnis ausgegangen, da im wissenschaftlichen Realismus an der Objektivitätsprämisse festgehalten wird. Zudem wird die Existenzbehauptung auf theoretische Annahmen gerichtet und nicht auf die Wirklichkeit empirisch direkt zugänglicher Zustände (Carrier, 1995a).

1.5.2 Methodisches Vorgehen

Methodenkombination

Der Entscheid für ein methodisches Vorgehen hängt nicht von der wissenschaftstheoretischen Orientierung ab (Hunt, 1991; Deshpandé, 1983). Für den empirischen Teil dieser Arbeit wird ein Methoden-Mix aus qualitativen und quantitativen Methoden gewählt. Damit wird der Erkenntnisgewinn erhöht, die spezifischen Stärken der Methoden berücksichtigt und dem Ziel Rechnung getragen, das Problem aus verschiedenen Perspektiven zu betrachten (vgl. Tab. 1-3; Mayring, 2002). Quantitative und qualitative Methoden sind nicht prinzipiell unvereinbare Untersuchungsverfahren, sondern als sich ergänzende Alternativen im Feld empirischer Forschung zu begreifen (Lamnek, 2005). Der Methodenmix verhindert zudem eine einseitige Fixierung auf inferenzstatistische Methoden des quantitativen Paradigmas (Mayring, 2002; Tomczak, 1992; Arndt, 1985).

Tabelle 1-3: Qualitative versus quantitative Forschung

	Qualitative Forschung	Quantitative Forschung
Art der Forschung	Explorativ / Interpretativ	Deskriptiv / Statistisch
Annahmen über Realität	Subjektiv konstruierte Realität, kontextfreie Untersuchung unmöglich	Objektiv gegebene Realität, kontextfreie Untersuchung möglich
Art der Fragestellung	Offen	Geschlossen
Art der Datengewinnung	Fallstudie, offenes Interview (unstrukturierte Verfahren)	Fragebogen, standardisiertes Interview (stand. Verfahren)
Anzahl der Befragten	Wenige	Viele
Analyse	Subjektiv	Statistisch
Qualifikation Interviewer	Hohe Anforderungen	Tiefe Anforderungen
Forschungsziel	Verstehen, Erklären, Gestalten	Erklären, Gestalten

Quelle: Eigene Darstellung in Anlehnung an Mayring, 2002 und Dutka, 1994

Hauptpostulate quantitativer Forschung sind Repräsentativität und intersubjektive Nachvollziehbarkeit. Bevor bei einer quantitativen Untersuchung die Datenerhebung beginnt, werden Entscheidungen über Theorie, Hypothesen, Erhebungsinstrument und Datenanalysemethoden getroffen (Stier, 1999). Das „reine“ quantitative Paradigma des kritischen Rationalismus erlaubt wegen des Falsifikationsprinzips nur deduktives, nomologisches Schliessen (Popper, 1973). Es ist ein positivistischer Forschungsansatz, bei dem die Überprüfung von generalisierbaren Hypothesen und Theorien im Vordergrund steht (Kroeber-Riel und Weinberg, 2003).

Qualitative Forschung geht von der konstruktivistischen Annahme aus, dass Realität subjektiv erzeugt ist. Sie interessiert sich primär für das „Wie“ der Zusammenhänge und deren innere Struktur, vor allem aus Sicht der jeweils Betroffenen (Bogdan und Taylor, 1975). Qualitative Forschungsmethoden sind wertvoll für die explorative Forschung, d.h. zur Entwicklung und Generierung innovati-

ver Theorieansätze (Deshpandé, 1983). Sie können im Sinne eines *empirischen Induktivismus* neue Hypothesen in den Forschungsprozess einbringen (Tomczak, 1992, S. 77).

Da in der vorliegenden Arbeit wenig bekannt ist über Hintergründe und Theorien zu Bestrafung und Anreiz-Massnahmen im Airline-Bereich, eignet sich für diesen Bereich die qualitative Forschung. Das in Kapitel 2.6 aufgeführte provisorische Wirkungsmodell wird mit der qualitativen Empirie zum zentralen Forschungsmodell erweitert und angepasst. Die dazugehörigen Hypothesen werden in Kapitel 4.2 vorgestellt und dienen als Grundlage zur konfirmativ-quantitativen Empirie.

Literaturanalyse

Ausgehend von der Literaturanalyse zum Thema *Kundensteuerungsmassnahmen mit Push- oder Pull-Massnahmen* (vgl. Kap. 2.1.2) und verwandten Konstrukten wie kritischen Ereignissen, Kundenzufriedenheit und Erwartungs-Diskonfirmation (vgl. Kap. 2.1.3; 2.5) werden zur Erarbeitung des provisorischen Wirkungsmodells verschiedene theoretische Perspektiven eingenommen (deduktive, d.h. theoriegeleitete Vorgehensweise). Zur Konzeptionalisierung des zu untersuchenden Phänomens der Bestrafung bietet es sich an, neben Theorien der Psychologie auch Theorien der Betriebswirtschaft und der Soziologie miteinander zu vereinen (vgl. Kap. 2.3; 2.4). Die Rechtswissenschaft liefert Einsichten zur Notwendigkeit, zum Sinn und zur Funktion von Strafen (vgl. Kap. 2.2). Diese von Feyerabend (1965) propagierte pluralistische Theorienauswahl vermag Reichweite und Gehalt realitätsorientierter Forschungsergebnisse zu erweitern (Ulrich und Hill, 1979).

Auf Basis des Literaturstudiums wird eine Ausgangsmenge an Einflussfaktoren für die theoretisch vermuteten Faktoren des Konstrukts *Bestrafung* gebildet (Homburg und Giering, 1996). Zudem werden das provisorische Wirkungsmodell und provisorische Hypothesen erarbeitet (vgl. Kap. 2.6).

Qualitative Methodik

Die Konzeptualisierung des Konstrukts *Kundenbestrafung* und die Konkretisierung des *a priori* zentralen Forschungsmodells werden mit verschiedenen Methoden der qualitativen Sozialforschung vorgenommen (exploratorische Empirie). Im Vordergrund stehen die Airline-Branche und die damit verbundenen Kundensteuerungsmassnahmen. Wichtig sind dabei diejenigen Strafmechanismen und Anreizsysteme, die Auswirkungen auf die Verhaltensabsichten von Passagieren haben (vgl. Kap. 3).

In einem ersten Schritt wurde eine E-Mail-Umfrage durchgeführt, um erste Erkenntnisse zu Strafen im Airline-Bereich zu gewinnen. Mit Expertengesprächen wurden die provisorischen Hypothesen aus der Literaturanalyse in einem zweiten Schritt validiert und Szenarien zur Fragebogenentwicklung der quantitativen Untersuchung entwickelt. Um die Kundenwahrnehmung direkt am „Objekt Passagier“ zu untersuchen und damit Verzerrungen beim Antwortverhalten in Interviews und Fragebogen zu unterbinden, drängte sich eine Beobachtung am Check-In auf. Mit der Beobachtung konnten einzelne Konstrukte und damit verbundene Hypothesen des zentralen Forschungsmodells untersucht werden. Mit der Inhaltsanalyse verschiedener Airline Websites wurde als letzter Schritt

der qualitativen Empirie untersucht, wie Airlines das Internet als Kanal zur Kundensteuerung von Passagieren nutzen können.

Quantitative Methodik

Für die quantitative Empirie wurden Probanden mit einem Experiment am Flughafen Zürich und mit einem Online-Experiment schriftlich befragt. Zielsetzung war die Überprüfung der Hypothesen und des zentralen Forschungsmodells (vgl. Kap. 4.2). Die Auswirkung verschiedener Einflussfaktoren wurde aus Airline-Perspektive auf die kognitive sowie die affektive Reaktion und die Verhaltensabsichten von Kunden aufgezeigt. Im Zentrum standen die Kundensteuerungsmassnahmen *Bestrafung* und *Anreiz-Massnahmen* sowie deren Kommunikation.

Die Hypothesenprüfung wird mit entscheidungsstatistischen Prüfverfahren und der Varianzanalyse durchgeführt; die Kriterien zur Auswahl der Methode werden in Kapitel 5 erläutert. Die quantitative empirische Untersuchung ist als Realitätsprüfung (konfirmative Empirie) zu verstehen und nicht als umfassende Theoriebestätigung (Bollen, 1989).

1.5.3 Forschungsprozess und Aufbau der Arbeit

Der Forschungsprozess kann basierend auf den oben ausgeführten Methoden und der Literaturanalyse in die Phasen *Beschreiben*, *Erklären* und *Gestalten* gegliedert werden. In Abbildung 1-2 werden der Fokus, die verwendeten Methoden und die Perspektive (Sicht) aufgeführt. Gleichzeitig widerspiegelt die Abbildung die Gliederung der Arbeit, die nachfolgend erläutert wird:

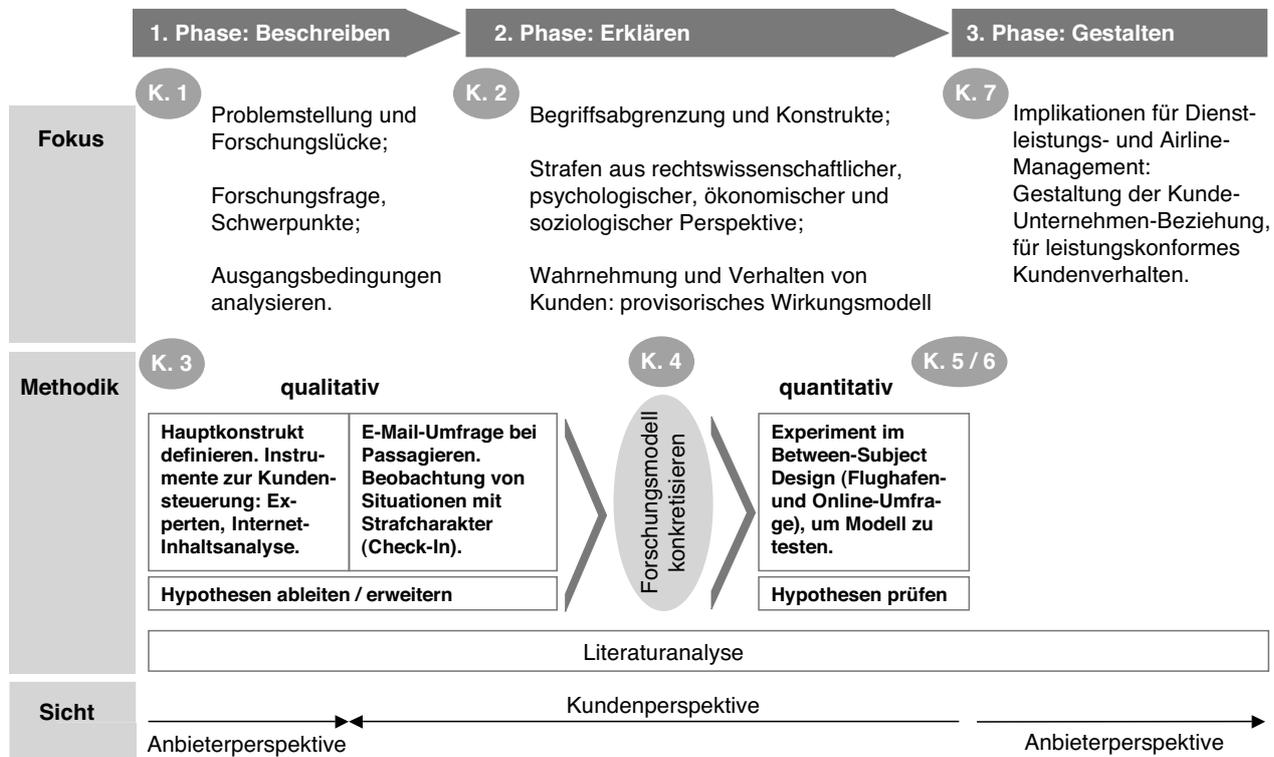
Kapitel 1: Einführung in die Problemstellung und Einordnung in den aktuellen Stand der Forschung. Ziele und Schwerpunkte werden vorgestellt und die wissenschaftstheoretische Positionierung sowie das Forschungsvorgehen als Kombination aus qualitativen und quantitativen Methoden erläutert.

Kapitel 2: Nachdem die relevanten Begriffe im Dienstleistungsbereich eingeführt sind, werden die Massnahmen zur Kundensteuerung sowie die Verhaltensabsichten von Kunden vorgestellt. Nach dieser Begriffsabgrenzung wird der rechtswissenschaftliche Hintergrund zu Strafen und verschiedenen Straftheorien erläutert. (Sozial-) Psychologische Modelle sowie Einsichten der Betriebswirtschaft und (Gruppen-) Soziologie als auch Erkenntnisse aus dem Dienstleistungsmarketing helfen, das provisorische Wirkungsmodell (und die provisorischen Hypothesen) aus Kunden- und Anbieterperspektive zu erstellen. Konstrukte, die mit Kundensteuerungsmassnahmen zusammenhängen und im Wirkungsmodell abgegrenzt werden müssen, sind ebenfalls erläutert.

Kapitel 3: Im qualitativen Teil der Arbeit werden die relevanten betriebswirtschaftlichen Grundlagen zur Airline-Branche, insbesondere *kritische Ereignisse* und *Kundensteuerungsmassnahmen*, vorgestellt. Anschliessend werden der Untersuchungsfokus und das Vorgehen der qualitativen Empirie erläutert. Aufbauend auf der Experten- und Kundenumfrage wird das Hauptkonstrukt *Bestrafung* aus Anbieter- und Kundenperspektive näher betrachtet. Nachdem eine erweiterte Dimensionsanalyse alle Aspekte der Kundenbestrafung geklärt hat, wird nachfrageseitig (Kundenperspektive)

eine Beobachtung am Check-In durchgeführt. Um die Kundensteuerung über das Internet untersuchen zu können, folgt eine Inhaltsanalyse der Airline Websites als Vorbereitung zur Fragebogenentwicklung der quantitativen Empirie.

Abbildung 1-2: Forschungsprozess und methodisches Vorgehen



Legende: K. = Kapitel

Quelle: Eigene Darstellung in Anlehnung an Schmidt, 2004

Kapitel 4: Basierend auf dem theoretischen, provisorischen Wirkungsmodell und den Ergebnissen der qualitativen Empirie werden das zentrale Forschungsmodell und die zu testenden Hypothesen formuliert. Das zentrale Forschungsmodell wird mit dem Flughafen-Experiment und dem Online-Experiment in zwei Schritten quantitativ getestet.

Kapitel 5: Nachdem die Grundlagen zum Experiment erarbeitet wurden, wird die Fragebogenentwicklung durchgeführt: Operationalisierung der unabhängigen Variablen, Skalenkonfiguration, Strukturierung des Fragebogens und Pretests. Die durchgeführte Datenerhebung und die Programme zur statistischen Auswertung werden für beide Experimente vorgestellt.

Kapitel 6: Nach der deskriptiven Datenanalyse werden die Ergebnisse des Hypothesentests vom Flughafen-Experiment einzeln vorgestellt. Anschliessend werden Anpassungen am zentralen Forschungsmodell vorgenommen. Das Testen der Modellanpassungen, die Replikation der zentralen Hypothesen und die Prüfung neuer Hypothesen werden mit einem zweiten, web-gestützten Experiment durchgeführt. Das Kapitel schliesst mit den empirischen Grenzen der Arbeit und den Folgerungen in Bezug auf die Forschungsfragen und das Forschungsmodell.

Kapitel 7: Es werden Implikationen zur Beantwortung der Forschungsfragen aus der empirischen Arbeit sowie der theoretischen Fundierung abgeleitet. Dabei werden Ansatzpunkte für das Management der Kundensteuerung dargestellt, wobei die Unternehmensziele berücksichtigt werden. Diese Schlussbetrachtung reflektiert den wissenschaftlichen und praktischen Beitrag der Arbeit. Die Forschungsergebnisse werden zusammengefasst und kritisch gewürdigt. Anschliessend werden Ansätze für die weitere Forschung formuliert.

1.6 Zusammenfassung

Im Serviceprozess kann Kundenverhalten leistungskonform gesteuert werden, indem Richtlinien mit Strafandrohung oder Anreiz-Massnahmen als Optionen eingeführt werden. Verschiedene Autoren weisen darauf hin, dass Kundenstrafen als Folge nicht eingehaltener Richtlinien in verschiedenen Dienstleistungsindustrien an Bedeutung gewinnen. Die Kommunikation dieser Kundensteuerung und der Preis einer Strafe resp. Option beeinflussen die Fairnesswahrnehmung von Kunden und ihre Verhaltensabsichten.

Die praktische Forschungslücke zeigt, dass in der Airline-Branche Kundensteuerungsmassnahmen nicht optimal umgesetzt sind. Gerade in der Airline-Branche besteht eine hohe Anfälligkeit für kritische Ereignisse, die als Strafe wahrgenommen werden können. Zudem gehören Bestrafungen heute zum Industriestandard in der Airline-Branche.

Das Forschungsziel der vorliegenden Arbeit ist die Untersuchung des Einflusses von Kundenstrafen auf die Verhaltensabsichten von Flugreisenden. Unter Einbezug juristischer, ökonomischer, psychologischer und soziologischer Forschungserkenntnisse wird das provisorische Wirkungsmodell entwickelt. Mit der Kombination von qualitativen und quantitativen Methoden werden das provisorische Wirkungsmodell vertieft, Hypothesen generiert und das zentrale Forschungsmodell mit zwei Experimenten getestet.

2 Theoretisch-konzeptionelle Grundlagen

Es gibt keine Theorie, die Kundenreaktionen und -verhalten auf Bestrafung durch Unternehmen umfassend erklärt. Daher wird die Thematik in den theoretisch-konzeptionellen Grundlagen aus verschiedenen wissenschaftlichen Perspektiven untersucht, nachdem die wichtigsten Begriffe bezüglich Kundensteuerungsmassnahmen im Dienstleistungsbereich definiert wurden (Kap. 2.1). Zuerst werden anhand einer juristischen Betrachtung Sinn und Funktion von Strafen im öffentlichen sowie im privaten Recht erläutert (Kap. 2.2). Schliesslich werden aus (sozial-) psychologischer Sicht die Hintergründe zur Wahrnehmung und zu den Entscheidungsprozessen von Konsumenten erklärt (Kap. 2.3). Basierend auf diesen Erkenntnissen wird das provisorische Wirkungsmodell entwickelt. Eine Betrachtung aus ökonomischer und soziologischer Sicht liefert Einsichten zur Bestrafungshärte, zum Verhältnis zwischen Individuen und Organisationen sowie zum Einfluss von Gruppen auf Strafsituationen (Kap. 2.4).

Die theoretisch-konzeptionellen Grundlagen werden mit einer Skizze des vorläufigen Erklärungs- und Wirkungsmodells abgeschlossen (Kap. 2.6). Hierzu werden vorerst die verwandten Konstrukte vorgestellt, wobei Erkenntnisse der Dienstleistungsmarketing- und Käuferverhaltensforschung zur affektiven Reaktion von Kunden bei Bestrafungen einfließen. (Kap. 2.5).

2.1 Begriffsabgrenzung

In diesem Kapitel werden zuerst die zentralen Begriffe *Dienstleistungsqualität* und die damit einhergehende *Dienstleistungsmarke (Image)* vorgestellt (Kap. 2.1.1). Die Dienstleistungsmarke wird im Forschungsmodell als (*ex ante*) Einflussfaktor auf die Kundensteuerung untersucht. Es wird zudem angenommen, dass die Kundensteuerungsmassnahmen eine Auswirkung auf die Markenwahrnehmung (*ex post*) haben. Anschliessend wird das Forschungsobjekt dieser Arbeit, die *Massnahmen zur Kundensteuerung*, für den Dienstleistungsbereich abgegrenzt (Kap. 2.1.2). Zudem wird in Kapitel 2.1.2 das Konstrukt *Bestrafung* definiert und den *Anreiz-Optionen* gegenübergestellt. Ein mit Strafen verwandtes Konstrukt sind die *Dienstleistungszwischenfälle*, die in Kapitel 2.1.3 erläutert werden. Dienstleistungszwischenfälle werden in dieser Arbeit nicht empirisch untersucht. Als Folge einer Bestrafung oder eines Dienstleistungszwischenfalles sind die *Verhaltensabsichten* von Kunden von Interesse (Kap. 2.1.4). Gemäss Forschungsmodell werden in der vorliegenden Untersuchung die Auswirkungen verschiedener Einflussfaktoren auf die Verhaltensabsichten von Kunden untersucht.

Grundlage der Entwicklung des Dienstleistungsmarketings war die Überzeugung von Forschern, dass Dienstleistungen im Vergleich zu tangiblen Gütern markante Unterschiede aufweisen und somit auch das Marketing spezifisch zu gestalten ist (Mattila und Patterson, 2004). Obwohl Shostack (1977) argumentiert, dass es wenig reine Produkte oder Dienstleistungen gibt, werden Dienstleistungen meist über den Vergleich mit Produkten definiert. Jedes Produkt weist bspw. einen Dienstleistungsteil auf (Kotler, 2006). Trotzdem lässt sich von einem Produkt der Dienstleistungsanteil

gesondert betrachten (Gabbott und Hogg, 1994). Es wird demnach anerkannt, dass Dienstleistungen einige charakteristische Merkmale aufweisen: *Intangibilität*, *Inseparabilität*, *Heterogenität*, *Vergänglichkeit* und *Eigentum* (vgl. A.1).

2.1.1 Dienstleistungsqualität und -marke

Die Interaktion zwischen dem Mitarbeiter einer Firma und dem Kunden wird als Dienstleistungsbegegnung (*Service Encounter*) oder *Moment of Truth* (Carlzon, 1987) bezeichnet. Surprenant und Solomon (1987) definieren die Dienstleistungsbegegnung als die zweiteilige (*dyadic*) Interaktion zwischen Kunden und Dienstleistungsunternehmen, wobei jede Partei ihre Rolle spielt. Breiter wird die Dienstleistungsbegegnung von Shostack (1985) definiert: „Eine Zeitperiode, während der ein Kunde mit einem Mitarbeiter interagiert“. Während der Dienstleistungsbegegnung ist die *Wahrnehmung* des Kunden eine entscheidende Komponente in der Evaluation der von der Firma bereitgestellten Dienstleistungsqualität (Barnes et al., 1999; Bitner et al., 1990). Die Dienstleistungsbegegnung spielt daher eine wichtige Rolle bei der Kundenzufriedenheit: je mehr Kunden Überraschung, Freude und andere positive Emotionen erfahren, umso höher wird der Zufriedenheitslevel sein (Oliver, 1997). Während einer Flugreise bestehen viele *Moments of Truth*, da es sich bei einer Flugreise um eine Dienstleistung mit hohem Interaktionsanteil handelt (*high contact service*). Zudem ist eine Flugreise eingebettet in weitere Dienstleistungsprodukte (Transport zum Flughafen, Dienstleistungen am Gate etc.).

Um das Konstrukt *Dienstleistungsqualität* erfassen zu können, haben Parasuraman et al. (1991) aus zehn Faktoren fünf Kategorien zur Dienstleistungsqualität erstellt (*Tangibles*, *Responsiveness*, *Reliability*, *Assurance* und *Empathy*). Grönroos (1988) hat später noch einen sechsten Faktor *Recovery* hinzugefügt (vgl. Kap. 2.1.3). Für Parasuraman et al. (1991) besteht die Dienstleistungsqualität aus der Differenz zwischen dem, was der Kunde erwartet, und dem, was der Kunde erhält (vgl. auch Kap. 2.5.1). Zur Evaluation der Dienstleistungsqualität muss die erwartete Qualität der wahrgenommenen Qualität gegenübergestellt werden (Bieger, 2003; Zeithaml, 2000; Zeithaml und Bitner, 2000). Diese Differenz wird beeinflusst vom Management einer Unternehmung: reale Kundenerwartungen und vom Management wahrgenommene Kundenerwartungen; die Spezifikation der Dienstleistungsqualität und die gelieferte Dienstleistung. Ein neueres Modell zur Bewertung von Dienstleistungsqualität wurde von Johnson et al. (1995) entwickelt, wobei aus fünfzehn Qualitätsaspekten drei Faktoren herausgearbeitet wurden:

1. *Hygiene Factors*: werden vom Kunden erwartet;
2. *Enhancing Factors*: führen zur Kundenzufriedenheit. Bei mangelhafter Lieferung entsteht aber keine Unzufriedenheit;
3. *Dual Threshold Factors*: Bei mangelhafter Lieferung dieser Faktoren entsteht Kundenunzufriedenheit, während Lieferung über einen spezifischen Level in Zufriedenheit mündet.

Die Abgrenzung zwischen dem Konstrukt Dienstleistungsqualität und dem Konstrukt *Kundenzufriedenheit* ist fließend. Zwar korrelieren diese beiden Konstrukte meist (hohe Dienstleistungsqualität impliziert in der oben dargestellten Definition hohe Kundenzufriedenheit), aber je nach Perspektive (Kunde oder Unternehmen) werden die Konstrukte verschieden erfasst (vgl. Definition von Kundenzufriedenheit in Kap. 2.5.2). Eine gute Abgrenzung, die für die vorliegende Arbeit verwendet wird, haben Rust und Oliver (1994) vorgenommen. Sie unterscheiden zwischen der Qualitätswahrnehmung, welche keine *Experience* mit der Dienstleistung oder dem Anbieter voraussetzt, und der Kundenzufriedenheit, welche *purely experimental* ist: Kunden können bezüglich einer Marke wohl geformte Qualitätswahrnehmung aufweisen, aber ohne Produkterfahrung keine Kundenzufriedenheitserfahrung haben.

Im Gegensatz zu Produkten kann eine Dienstleistung nicht visualisiert werden. Unternehmen sind deshalb gefordert, immaterielle Leistungskomponenten physisch darzustellen und diese durch Symbole des Markenzeichens zu unterstützen (Meffert und Bruhn, 2000). Analog zu Produktmarken ist bei *Dienstleistungsmarken* das Versprechen eines einheitlichen Qualitätsniveaus entscheidend (Stauss, 1998). Dies ist jedoch aufgrund der Heterogenität externer Faktoren für das Dienstleistungsunternehmen schwierig zu erreichen. Dienstleistungsmarken beschränken sich aufgrund der Intangibilität der Leistung vorwiegend auf Unternehmensnamen (Kreuzpaintner, 2003). Dem Kunden soll Sicherheit und Vertrauen vermittelt werden.

Ein Image entsteht mit der Marken-Wahrnehmung des Kunden (vgl. Kap. 2.3.1): "*The brand as a concept is always an image*" (Grönross, 2000, S. 285). Marken helfen den Konsumenten, Entscheidungen zu routinisieren, Komplexität zu reduzieren und damit den Konsumenten in dessen Entscheidungsfindung zu entlasten (Drenger, 2006).

Es gibt keine einheitliche Definition zum Begriff *Marke* oder *Image* (Keller, 1993). Verschiedene Autoren betonen die Mehrdimensionalität des Konstrukts und bauen die Definition auf verschiedenen psychologischen Wahrnehmungskonzepten auf (Drenger, 2006). Ein Ansatz postuliert bspw., dass das Konstrukt *Image* ähnliche Merkmale wie das Konstrukt *Einstellung* besitzt. Gemäss diesem Ansatz können daher beide Begriffe synonym verwendet werden (vgl. Kap. 2.3.1 zu *Einstellungen*; Kroeber-Riel und Weinberg, 2003). *Image* wird in dieser Arbeit in Anlehnung an Trommsdorff (2003) und Meffert (2000) als ein mehrdimensionales, mit subjektiven Wertungen versehenes, komplexes psychisches Konstrukt verstanden. Das *Image* einer Firma wird bei Konsumenten gemäss dem in der vorliegenden Arbeit verwendeten Basismodell zur Wahrnehmung erzeugt (vgl. Kap. 2.3.1). Die Begriffe *Marke* und *Image* werden in dieser Arbeit synonym verwendet.

2.1.2 Massnahmen zur Kundensteuerung

Unternehmen können ihre Kunden neben der Implementierung von Richtlinien und Strafen auch mit Anreizmechanismen wie *Bonusprogrammen* oder *Cross- und Upselling* steuern. Wie bereits in Kapitel 1.4.1 erläutert, werden die *Bestrafung* und *Anreizmechanismen* als definatorische „Spiegelbil-

der“ betrachtet. Diese als *Push-* oder *Pull-Massnahmen* bezeichneten Kundensteuerungsmassnahmen können singular, sequenziell oder simultan eingesetzt werden (Lauer, 2004). Der Unterschied der Massnahmen liegt in der Intensität der Kundensteuerung und den Auswirkungen auf die Kundenwahrnehmung sowie den Verhaltensabsichten. Kundensteuerungsmassnahmen werden im vorliegenden Forschungsmodell als Einflussfaktoren auf die Kundenwahrnehmung und auf die Verhaltensabsichten identifiziert.

Strafen als Push-Massnahmen

Im Dienstleistungssektor hat sich die Bestrafung von Kunden durchgesetzt, weil im Gegensatz zum produzierenden Sektor das Dienstleistungsprodukt intangibel und nicht lagerfähig ist. Einmal bereitgestellt, kostet das offerierte Dienstleistungsprodukt und muss konsumiert werden, um einen Deckungsbeitrag zu leisten. Nicht konsumierte Dienstleistungen sind reine Kostenträger und hätten vom Dienstleistungsunternehmen anderweitig verkauft werden können (vgl. Kap. 1.1). Die Bestrafung ist in dieser Hinsicht eine *Kompensation* für die verlorene Opportunität, wenn ein Kunde seiner geäusserten Konsumabsicht nicht nachkommt. Ziel ist es, das Kundenverhalten während des Serviceprozesses durch Androhung von Richtlinien, deren Nichteinhaltung Strafen nach sich ziehen, zu steuern (*Prävention*).

Die Kundensteuerung über Richtlinien und Strafen wird als *Push-Massnahme* bezeichnet: Kunden werden gezwungen, ihre „gewohnten“ Verhaltensmuster zu ändern (vgl. Schögel und Schulden, 2006; Bansal et al., 2005). Dadurch wird der wahrgenommene Kundenvorteil im Gegensatz zu *Pull-Massnahmen* reduziert. Der *Kundenvorteil* entspricht dem vom Kunden wahrgenommenen Nutzenszuwachs durch eine Leistung (Belz und Bieger, 2006).

Der Brockhaus definiert den Begriff der Strafe als „ein Übel, das jemand einem anderen mit Absicht zufügt, weil dieser eine missbilligte Handlung begangen hat. Das Ziel der Strafe ist, die Wahrscheinlichkeit des Auftretens der missbilligten Handlung herabzusetzen und das Verhalten eines Menschen zu verändern“ (o.V., 2006a, S. 397). Angewandt auf den Kontext *Unternehmen-Kunde* bedeutet dies, dass mit Strafen das Verhalten von Kunden verändert werden soll. Kundenbestrafung kann entweder monetär (*Penalty*) oder nicht monetär sein (*Punitive Actions*). Aus Sicht des Kunden spielen dabei Wahrnehmungsprozesse eine wesentliche Rolle (vgl. Kap. 2.3). In der Forschung zu Bestrafungen im Dienstleistungsbereich wird die monetäre Bestrafung wie folgt definiert:

„*Monetäre Bestrafung ist die Auferlegung einer Gebühr an Kunden, die es versäumen, die originäre Einkaufsvereinbarung (Richtlinie) zu erfüllen*“ (McCarthy und Fram, 2000; Fram, 1997).

In der Beziehung zwischen Unternehmen und Kunde werden meistens Gebühren resp. Geldstrafen als ökonomische Form der Bestrafung angewendet, um nicht erwünschtes Kundenverhalten abzuwenden. Strafen sind für viele Dienstleistungsunternehmen ein integraler Teil der Preisstrategie (Kim, 2006). Monetäre Bestrafungen haben somit verschiedene Ziele. Einerseits sollen Kunden von unerwünschtem Verhalten abgehalten werden, um Kosten zu sparen oder um fehlende Deckungs-

beiträge zu kompensieren (negativer Stimulus, vgl. Kap. 2.3.1). Andererseits können mit monetären Bestrafungen zusätzliche Umsätze generiert oder bei unerwünschtem Verhalten der ökonomische Verlust vermindert werden (McCarthy und Fram, 2000). Letztlich können potentielle Verluste verhindert werden, indem Unternehmen umsatzschwache oder unerwünschte Kundensegmente abschrecken (z.B. Strafgebühren im US-Bankwesen für zu tiefen Kontostand). Entweder soll die Bestrafung zusätzlichen Ressourcen- und administrativen Aufwand kompensieren, Transaktions-, Risiko- und Opportunitätskosten decken, die durch schädliches Kundenverhalten entstanden sind, oder Umsatz generieren (Kim und Smith, 2005). Die durchschnittliche Gebühr von Kreditkartenunternehmen für zu späte Einzahlung einer Rechnung ist bspw. von 13 USD im Jahr 1995 auf über 25 USD im 2006 angestiegen (Kim, 2006). Viele Dienstleistungen werden heute günstig und attraktiv offeriert, um durch Bestrafungszahlungen verpasstes Einkommen aufzufangen. Billigairlines können bspw. tiefe Preise anbieten und gleichzeitig hohe Änderungsgebühren für die Ticket-Neuausstellung verlangen.

Konzeptuell sind monetäre Kundenbestrafungen berechtigt, weil Kunden eine bestimmte Rolle innerhalb der Dienstleistungskette durch die Firma einnehmen (Bitner, 1995). Sinngemäß ist es schwierig zu argumentieren, dass Firmen kein Recht haben, monetäre Strafen zu erheben (vgl. Sorrell, 1994). *“Customers may always be right, but allowing them to follow their own preferences often increases a company’s costs while leaving opportunities to boost revenues untapped. Instead, customers must be guided [...]”* (Myers et al., 2004, S. 37). Beispiele für monetäre Bestrafung sind:

- Strafen für Rückzug resp. Änderung einer Buchung oder für zu späte Bezahlung einer Rechnung (Fram, 1997);
- Strafzahlung für Nichtbenützung einer Kreditkarte (Shure, 1996) oder von Ferngesprächsguthaben einer Telefongesellschaft (Van, 1999);
- Gebühr für Kosten der Lagerhaltung eines umgetauschten Gutes (Woodyard, 1997);
- Gebühr für eine nicht stornierte Restaurant-Reservierung (Hwang und Green, 1999);
- Bestrafung von Kunden, die ein Billett eines Transportunternehmens verlieren resp. falsch benutzen (o.V., 1999).

Obwohl monetäre Bestrafung im Dienstleistungsbereich die gebräuchlichste Bestrafungsart ist, gibt es nicht monetäre Bestrafungen wie bspw. Wechselbarrieren zur Konkurrenz aufgrund technologischer Restriktionen und Verträgen, Verweigerung eines Dienstes oder Aufhebung von Privilegien durch den Dienstleistungsanbieter (Tomczak und Dittrich, 1997). Für einen Überblick über die Dimensionen, die eine Kundenstrafe umfassen, sei auf Kapitel 3.4 verwiesen. Die Abgrenzung zwischen Strafen und Anreizen wird in Tabelle 2-1 vorgenommen.

Tabelle 2-1: Massnahmen zur Kundensteuerung

	Push-Massnahmen	Pull-Massnahmen
Kundenwahrnehmung	<ul style="list-style-type: none"> – Strafe (Folge einer missachteten Richtlinie); – Meist keine Wahl (verpasste Opportunität). 	<ul style="list-style-type: none"> – Nutzen (Option als Anreiz); – Wahlmöglichkeit (Opportunität).
Instrumente	<ul style="list-style-type: none"> – Regeln: Einkaufsvereinbarung, Wechselbarrieren (technologisch oder vertraglich) etc.; – Gebühren oder Geldstrafen. 	<ul style="list-style-type: none"> – Produktgestaltung, Distribution, Kommunikation; – Bonusprogramm, Cross- und Upselling.
Unternehmensziel	<ul style="list-style-type: none"> – Abschrecken unerwünschter Kundensegmente; – Verhalten steuern, Kosten senken; – Kurzfristig Umsatz generieren. 	<ul style="list-style-type: none"> – Neue Kundensegmente ansprechen (Differenzierung); – Verhalten steuern, Umsatz steigern; – Kundenbindung und langfristig Umsatz generieren.

Quelle: Eigene Darstellung

Im Durchschnitt können in Grossbritannien ca. 250'000 Passagiere pro Jahr ihren Flug nicht antreten, weil entweder das Flugzeug überbucht ist oder der Flug storniert wurde (o.V., 2002a). Handelt es sich hier um ein negatives kritisches Ereignis oder um eine Bestrafung des Kunden, weil er bspw. wegen seines Status beim Boarding abgewiesen wurde? Die Literatur zu Bestrafungen und kritischen Ereignissen im Dienstleistungsbereich liefert keine schlüssige Antwort, sondern weist auf das Konzept der nicht monetären Strafen hin. Ein zentrales Element der nicht monetären Bestrafung ist die Kundenwahrnehmung nach einem Ereignis, das durch die Verletzung einer Richtlinie entstanden ist (Kumar et al., 1998). Daraus folgt die in dieser Arbeit verwendete Definition zur Bestrafung im Dienstleistungsbereich, wobei nicht zwischen monetären oder nicht monetären Strafen unterschieden wird:

„Bestrafung ist ein Verhalten von Firmen, welches in der Kundenwahrnehmung negative Konsequenzen verursacht oder positive Konsequenzen verunmöglicht. Ausgangspunkt der Bestrafung ist eine vom Kunden nicht eingehaltene Einkaufsvereinbarung.“

Anreize als Pull-Massnahmen

Das Kundenverhalten kann mit Anreiz-Massnahmen wie Anpassungen des Produkts (Leistungsumfang, Qualität, Preisgestaltung, Wechselbarrieren etc.), der Distribution (Internet, Kooperationen etc.) oder mit der Kommunikation (Werbung, PR etc.) gesteuert werden. Anreiz-Massnahmen werden als *Pull-Massnahmen* bezeichnet, weil dem Kunden ein höherer Nutzen bei verändertem Verhalten versprochen wird. Die Vorteile und die Wahrnehmbarkeit des „neuen Verhaltens“ werden erhöht (vgl. Schögel und Schulten, 2006; Bansal et al., 2005). Daraus folgen ein höherer Kundenvorteil und langfristiges Umsatzwachstum, da neue Kundensegmente angesprochen werden.

In dieser Arbeit werden unter Anreiz-Massnahmen (Optionen) das *Bonusprogramm* sowie *Cross- und Upselling* verstanden. Sie werden als das logische „Gegenkonzept“ zu Bestrafungen empirisch untersucht. Mit der Anreizsteuerung können wie bei Richtlinien die Kundenbetreuungskosten gesenkt werden (Lauer, 2004). Das Ziel der Anreizsteuerung ist die Steigerung des Kundenwerts durch erhöhten Umsatz und Kundenbindung. Der Kundenwert wird in dieser Arbeit definiert als der „aktuelle und künftige Beitrag von Kunden zum Erfolg des Unternehmens“ (Belz und Bieger, 2006). Er kann als Barwert aller Ein- und Auszahlungen pro Kunde berechnet werden (Lauer, 2004). Kundenwerte entstehen durch Kundenvorteile (Belz und Bieger, 2006).

Aus der Perspektive einer wertorientierten Kundensteuerung besteht das Ziel der Pull-Massnahmen, das Kundenverhalten potentieller und bestehender Kunden durch einen für den Kunden wahrnehmbaren Zusatznutzen in Einklang mit den Unternehmenszielen zu bringen (vgl. Schögel und Schulten, 2006).

Sogenannte *Bonusprogramme* sind als Angebot von Unternehmen an Kunden zu verstehen, die den Kunden für bestimmte Verhaltensweisen belohnen (Lauer, 2004). Neben der Kundensteuerungsfunktion durch eine gezielte Incentivierungs- und Einlösemechanik spielt die Erlangung eines besonderen Kundenstatus mit Zusatzservices für einen definierten Zeitraum eine wesentliche Rolle. Das Bonusprogramm muss dem Kunden klar und verständlich kommuniziert werden (Tomczak und Dittrich, 1997). Ein wichtiger Funktionsbestandteil von Kundenbindungsprogrammen ist der Internetauftritt als zentrale Kommunikationskomponente (Stahl, 2000). Bei Bonusprogrammen ist die Reaktanz von Kunden tiefer, da sie „nicht auf der tatsächlichen oder impliziten Androhung von *Strafe*, wie bei der vertraglichen oder technologischen Bindung“ beruhen (Lauer, 2004, S. 14). Die Kundenbindung wird erhöht, *Cross-* und *Upselling* gefördert, das Image sowie das Weiterempfehlungsverhalten werden gesteigert und schliesslich die Kundenbetreuungskosten vermindert (vgl. Buser und Welte, 2006; Berger und Wentzell, 2001). Die Kundenbetreuungskosten werden vermindert, indem (a) kostensparendes Verhalten durch den Kunden über Bonuspunkte belohnt wird oder (b) Kundenwissen genutzt und damit die Erfolgswahrscheinlichkeit von Werbemassnahmen erhöht wird (Lauer, 2004).

Die Untersuchung von Stauss et al. (2005) bezüglich Kundenfrustration in Bonusprogrammen hat die Nachteile eines Bonusprogramms aufgezeigt. Folgende Barrieren können bei Kunden zur *Entstehung von Frust* (vgl. Kap. 2.5.2) in einem Bonusprogramm eines Transportunternehmens führen:

- Stuserlangung (*qualification barrier*);
- Unmöglichkeit, die Belohnung zu fordern (*inaccessibility barrier*);
- Wertlosigkeit der Belohnung (*worthlessness barrier*);
- Zusatzkosten oder Zusatzaufwand, um Belohnung einzufordern (*redemption barrier*).

Zusätzlich zu den oben genannten Frustereignissen haben die Probanden wegen des Bonusprogramms eine Verschlechterung der Beziehung zum Dienstleistungsunternehmen festgestellt. Sie fühlten sich diskriminiert wegen der besseren Behandlung einzelner Passagiere und fanden, dass die Firma falsche Prioritäten im Bonusprogramm setzt. Anstatt sich auf die Kerndienstleistung zu konzentrieren und diese zu verbessern, wurden gemäss Stauss et al. (2005) zu viele Ressourcen in das Bonusprogramm gesteckt. Interessanterweise hat sich mit dem Aufbau von Bonusprogrammen auch der Einsatz von Strafrichtlinien entwickelt. Es ist nicht unüblich, beide Strategien in einem Unternehmen gleichzeitig anzutreffen (Kim, 2003).

Dem Kunden können während des Kaufprozesses neue Leistungen (*Cross-Selling*) oder Zusatzleistungen (*Upselling*) angeboten werden. Der Zeitpunkt, wo der Kunde telefonisch oder elektronisch bereits mit dem Unternehmen in Verbindung steht, ist ideal zur Anreizsteuerung als Teil des *Customer Relationship Managements* (vgl. Wong und Sohal, 2003). *Cross-* und *Upselling* gegenüber Flugpassagieren bezieht sich dabei auf das Anbieten von Zusatzoptionen zum Flugticket (z.B. Sitzplatzreservation bei *Low Cost Carrier*) oder Dienstleistungen von Partnerunternehmen der Airline wie z.B. Reiseversicherungen, Hotelbuchungen oder Autovermietung (Fleischer, 2004). Erstaunlich ist, dass gemäss Expertenbefragung (vgl. Kap. 3.4) beim *Upselling* vor allem den Kunden mit den tiefsten *Economy*-Tarifen ein *Upgrade* in die *Business Class* verkauft werden kann.

2.1.3 Dienstleistungszwischenfall

Dienstleistungszwischenfälle sind definiert als kritische Ereignisse (*Critical Incidents*), die zur Erreichung eines generellen Ziels einer Aktivität beitragen oder dabei hinderlich sind (Flanagan, 1954). Im Kontext dieser Arbeit ist ein kritisches Ereignis eine Interaktion zwischen Kunde und Dienstleistungsunternehmen. Diese Interaktion kann besonders positiv oder negativ in Erinnerung bleiben (vgl. Ronan und Latham, 1974).

Um die Begriffe kritisches Ereignis sowie Bestrafung abzugrenzen, sind diese Konstrukte in Tabelle 2-2 aus Kunden- und Unternehmensperspektive aufgeführt. Für das Konstrukt *Diskonfirmation* wird auf Kapitel 2.5.1 verwiesen.

Kritische Ereignisse rufen nicht zwangsläufig negative Kundenreaktionen hervor. Erstens kann eine erfolgreiche Wiedergutmachung in einzelnen Fällen zu einer Verbesserung der Wahrnehmung von Dienstleistungsqualität führen (Chung-Herrera et al., 2004). Dies wird *Recovery Paradox* genannt (Andreassen, 2000). Zweitens gibt es positive kritische Ereignisse, wie bspw. unerwarteter Rabatt, welche die Kundenzufriedenheit, Loyalität und die Wiederkaufswahrscheinlichkeit erhöhen (Bamford und Xystouri, 2005; Wong und Sohal, 2003).

Obwohl negative kritische Ereignisse und Bestrafungen eng verwandt sind (eine Bestrafung könnte als kritisches Ereignis definiert werden), bestehen Unterschiede. Im Gegensatz zu Bestrafungen entstehen negative kritische Ereignisse aufgrund von Kundenerwartungen an die Dienstleistung.

Tabelle 2-2: Abgrenzung zwischen kritischen Ereignissen und Bestrafung

	Anbieterperspektive	Kundenperspektive
Kritische Ereignisse		
Positive Ereignisse:	– Optimale Prozesse;	– Positive Diskonfirmation;
Negative Ereignisse:	– Prozesse nicht normgemäss durchgeführt.	– Negative Diskonfirmation.
Bestrafung	<ul style="list-style-type: none"> – Missachtung von Richtlinien (Einkaufsvereinbarung); – Verhindern von nicht-leistungskonformem Verhalten. 	<ul style="list-style-type: none"> – Zahlung einer Gebühr; – Negative Konsequenz verursacht oder positive Konsequenz verunmöglicht.

Quelle: Eigene Darstellung

Werden die Erwartungen nicht erfüllt, spricht man von kritischen Ereignissen oder negativer Diskonfirmation (vgl. Kap. 2.5.1; Chung-Herrera et al., 2004). Ein weiteres Unterscheidungsmerkmal zwischen kritischen Ereignissen und Bestrafungen ist die Ursache des Fehlverhaltens. Negative kritische Ereignisse gründen im *Fehlverhalten der Unternehmung*. Bestrafungen gründen im *Fehlverhalten des Kunden* (vgl. Kap. 2.1.2).

Ein Dienstleistungszwischenfall (*Service Failure*) ist ein negatives kritisches Ereignis und eine Voraussetzung, damit Wiedergutmachung (*Service Recovery*) stattfinden kann (de Coverly et al., 2002). Der Kunde entscheidet, was ein Dienstleistungszwischenfall ist (Michel, 2001; Bitner et al., 1990). Langjährige resp. loyale Kunden verzeihen Fehler im Dienstleistungserstellungsprozess eher (Boulding et al., 1993). Chung-Herrera et al. (2004) haben basierend auf den drei *Failure Groups* von Bitner et al. (1990) negative kritische Ereignisse aufgrund einer Kundenbefragung bei Dienstleistungsanbietern den drei Kategorien zugeordnet:

- *Response to Service Delivery System Failures*: langes Anstehen, Fehlen von Komponenten der Dienstleistung (z.B. Reservation verloren), Fehler im Dienstleistungsprozess (z.B. Gepäckschäden) und Firmenrichtlinien (z.B. unklare Richtlinie);
- *Response to Implicit / Explicit Customer Requests*: Fehler bei Kunden mit Handicap (z.B. medizinische Versorgung), Vernachlässigung spezieller Kundenwünsche (z.B. vegetarisches Essen) und falsche Reaktion bei Kundenfehlern (z.B. Ticket verloren);
- *Unprompted / Unsolicited Employee Actions*: Fehlender Kundenrespekt, Anstand und Individualisierung (z.B. Kunden ignorieren), Fehler bezüglich kultureller Normen und Fehler wegen Inkompetenz des Mitarbeiters.

Der Fehlertyp, den Kunden als am ärgerlichsten empfanden, sind unfaire sowie unklare Firmenrichtlinien (Chung-Herrera et al., 2004).

2.1.4 Verhaltensabsichten

Der Forschungsbereich von Konsumentenverhalten ist breit: *“It is the study of the processes involved when individuals or groups select, purchase, use, or dispose of products, services, ideas, or experiences to satisfy needs and desires”* (Solomon, 2006, S. 7). Das Verhalten von Käufern wird bezeichnet als das beobachtbare *äußere* und das nicht beobachtbare *innere* Verhalten von Menschen beim Kauf und Konsum wirtschaftlicher Güter (Kroeber-Riel und Weinberg, 2003). Im vorliegenden provisorischen Wirkungsmodell werden die Verhaltensabsichten als abhängige Variablen aufgenommen, die aus der Kundenwahrnehmung einer Situation resultieren (vgl. Kap. 2.6).

Es wird bewusst keine Unterscheidung zwischen der Verhaltensabsicht (*Behavioral Intention*) und dem konkreten Verhalten gemacht. Kundenverhalten (*Customer Behavior*) ist eine direkte Konsequenz aus den Verhaltensabsichten. Konkretem Kundenverhalten ist messtechnisch schwieriger beizukommen, weil die Beobachtungsdauer nicht zu begrenzen ist. Ein tatsächliches Kundenverhalten kann auf Gründen basieren, die Jahre zurückliegen oder am Tag vorher auftraten. Aus diesem Grund kann die retrospektive Frage an einen Kunden, warum er sich genau so verhalten hat, verfälscht sein. Zudem betonen verschiedene Autoren, dass die Verhaltensabsichten ein ausgezeichneter Prediktor für tatsächliches Kundenverhalten sind (Cronin et al., 2000; Chang und Wildt, 1994; Bolton und Drew, 1991; Reichheld und Sasser, 1990).

Das tatsächlich resultierende Verhalten aus den Verhaltensabsichten ist entscheidend, weil es sich auf den finanziellen Erfolg eines Unternehmens auswirkt (Rudolf-Sipötz und Tomczak, 2001). Einige Forscher postulieren daher einen engen Zusammenhang zwischen den Verhaltensabsichten von Kunden und dem Unternehmenserfolg (vgl. Bougie et al., 2003; McNaughton et al., 2002; Cronin et al., 2000; Chang und Wildt, 1994).

Tabelle 2-3: Negative und positive Verhaltensabsichten

Negative Verhaltensabsichten	Positive Verhaltensabsichten
– Absicht, eine Beschwerde einzureichen;	– Absicht, Unternehmung zu loben;
– Negative Word-of-Mouth (WOM: Mund-zu-Mund-Propaganda);	– Weiterempfehlung, positive Word-of-Mouth;
– Illoyalität (kein Wiederkauf der Leistung oder Wechsel des Anbieters).	– Loyalität (Kundentreue, Wiederkauf).

Quelle: Eigene Darstellung in Anlehnung an Zeithaml et al., 1996

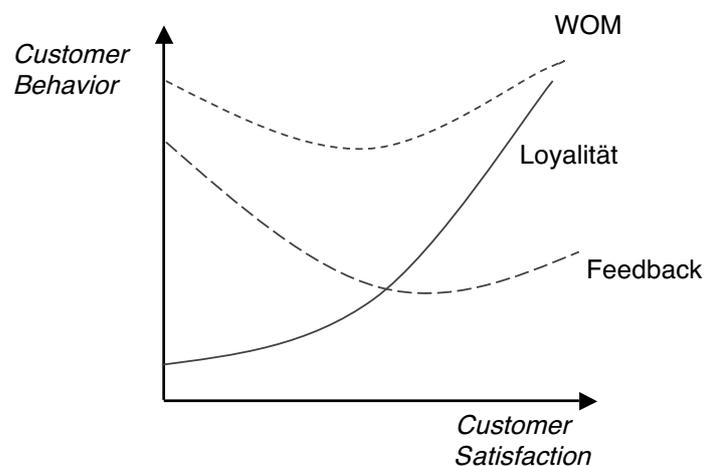
Die Verhaltensabsichten werden von Zeithaml et al. (1996) allgemein gegliedert als: Mund-zu-Mund-Propaganda (*positive Word-of-Mouth*), Weiterempfehlungen (*Recommendations*), Kundentreue (*Loyalty*), höhere Kauffrequenz (*Higher Consumer Expenditure*) und Bereitschaft zum Bezahlen von Preispremiums (*Willingness to Pay Price Premiums*). Im Kontext der vorliegenden For-

schungsfragen werden drei Verhaltensabsichten identifiziert (vgl. Tab. 2-3), als *negative Verhaltensabsichten* bei Strafen und *positive Verhaltensabsichten* bei Anreiz-Massnahmen (vgl. Solomon, 2006; White und Yu, 2005; Athanassopoulos, 2000; Andreassen, 1999; Söderlund, 1998; Zeithaml et al., 1996).

Diese Kundenabsichten können einzeln oder kombiniert auftreten. Solomon (2006) identifiziert neben den oben erwähnten Reaktionen auch die sogenannte *Third-Party Response*: Unzufriedene Kunden können rechtlich gegen das Dienstleistungsunternehmen vorgehen oder sich bei Konsumenten-schutzvereinen sowie Medien negativ über die Firma äussern (Andreassen, 2000; Singh, 1990b; Day et al., 1981).

Verhaltensabsichten werden u.a. durch das Konstrukt *Kundenzufriedenheit* ausgelöst (Athanassopoulos, 2000; Zeithaml et al., 1996). Kundenzufriedenheit wiederum ist ein Konstrukt, welches schwierig zu fassen ist (vgl. Kap. 2.5.2). Söderlund (1998) hat den Einfluss von verschiedenen Zufriedenheitslevels auf *Word-of-Mouth* (WOM), Loyalität und Feedback als Verhaltensabsichten gemessen. Dabei hat er herausgefunden, dass je nach Zufriedenheitslevel von Kunden für die drei erwähnten Verhaltensabsichten verschiedene Entwicklungen bestehen (vgl. Abb. 2-1).

Abbildung 2-1: Kundenverhalten und Kundenzufriedenheit



Quelle: Söderlund, 1998

Während die Loyalität mit dem Zufriedenheitslevel zunimmt, sinken sowohl WOM als auch die Feedback-Rate (Anzahl Beschwerden) bis zu einem bestimmten Zufriedenheitslevel. Danach steigen beide Kurven wieder an und erreichen ihr Maximum bei extrem hoher Zufriedenheit (Kompliment an Unternehmung). Das Feedback hat jedoch im Ausgangspunkt den Maximalwert erreicht und widerspiegelt die starke Reaktion auf negative kritische Ereignisse sowie Bestrafungen.

In der vorliegenden empirischen Untersuchung stehen negative Verhaltensabsichten im Vordergrund, da Bestrafungen das Hauptforschungsobjekt sind. Negative Emotionen (nach einer Bestrafung) setzen bei Individuen ein Signal frei, so dass eine Aktion eingeleitet werden muss (Loewen-

stein et al., 2001; Taylor, 1991). Positive Emotionen setzen kein *Aktionssignal* frei. Negative Ereignisse lösen zudem stärkere kognitive Evaluationen und Repräsentation aus (Peeters und Czapinski, 1990).

Loyalität

Das Konstrukt *Loyalität* ist rein semantisch betrachtet kein Verhalten. Obwohl in der Literatur Loyalität als eine eindimensionale Grösse definiert wird (*intention to purchase again from the supplier*), wird sie in dieser Arbeit als zweidimensionale Grösse aufgefasst und mit *Illoyalität* bezeichnet (kein Wiederkauf der Leistung oder Kauf bei einem Konkurrenzunternehmen). Diese Absichten hängen von der Kundenzufriedenheit und diese wiederum von der wahrgenommenen Dienstleistungsqualität ab (Söderlund, 1998). Von Keaveney (1995) wurden die *Switching Intentions* (Kauf bei einem Konkurrenzunternehmen) von illoyalen Kunden im Dienstleistungsbereich untersucht. Sechs Gründe haben zur erhöhten Neigung des Anbieterwechsels geführt: Preis, Unbequemlichkeit, Kerndienstleistungsfehler, Fehler in der Dienstleistungsbegegnung, attraktive kompetitive Angebote und ethische Probleme. Andere Forscher haben zusätzlich die Rolle von *Value for Money* betont (vgl. Athanassopoulos, 2000).

Der Zufriedenheitslevel des Kunden bestimmt die Loyalität (Zeithaml et al., 1996; Jones und Sasser, 1995). Die Zufriedenheit des Kunden wiederum hängt im Zusammenhang mit Bestrafungen davon ab, wie fair er sich behandelt fühlt, respektive wie die Gerechtigkeitswahrnehmung ist (McColl-Kennedy und Sparks, 2003). Dabei kann es vorkommen, dass selbst Kunden, die sehr zufrieden mit der Dienstleistung sind und sich fair behandelt fühlen, den Anbieter wechseln. Es gibt also keine absolute Loyalität; nur eine Loyalität gegenüber zwei oder drei verschiedenen Brands (Reichheld, 1996). Trotzdem setzen einige Forscher Kundenzufriedenheit als „Proxy“ Masseinheit für Loyalität ein (vgl. McColl-Kennedy und Sparks, 2003). Erhöhte Kundenzufriedenheit bedeutet aber nicht zwangsläufig erhöhte positive Kundenreaktionen (O'Malley, 1998).

Word-of-Mouth

Word-of-Mouth (WOM) wird definiert als das Ausmass, an dem Kunden ihre Freunde, Verwandte und Kollegen über ein Ereignis informieren. Dieses Ereignis hat eine hohe Zufriedenheit oder Unzufriedenheit ausgelöst (Söderlund, 1998). Weil WOM meistens durch Bekannte oder verwandte Personen erfolgt, erweist es sich für den Rezipienten als weitaus verlässlicher und vertrauensvoller im Vergleich zu bekannten Marketing-Instrumenten: Es wird angenommen, dass zwei Drittel aller Verkäufe im Konsumgüterbereich aufgrund von WOM erfolgen (Solomon, 2006).

Bei kritischen Ereignissen wirkt sich aber WOM gegenteilig aus: "*Consumers weigh negative Word-of-Mouth more heavily than they do positive comments*" (Solomon, 2006, S. 396). Fisk et al. (1990) und Hart et al. (1990) haben asymmetrische Effekte von positiven und negativen kritischen Ereignissen beobachtet: negative Ereignisse werden im Durchschnitt elf Verwandten / Freunden mitgeteilt, während positive Erlebnisse nur sechs Verwandten / Freunden erzählt werden.

Negative WOM kann durch verzerrte Kommunikation zwischen Konsumenten oder durch Veröffentlichung auf dem Internet zu schweren Reputationsschäden einer Unternehmung führen (Solomon, 2006).

Beschwerden

Bei einer Untersuchung des Office of Fair Trading (1999) stellte sich heraus, dass drei Viertel aller Kunden, die sich beschwerten und deren Problem erfolgreich von der Unternehmung gelöst wurde, wieder beim selben Unternehmen einkaufen. In den Fällen, wo mangelhaft oder nicht auf die Beschwerde eingegangen wurde, wird weniger als die Hälfte der Kunden bei der betreffenden Firma wieder einkaufen. Auch Forschungsergebnisse von Eccles und Durand (1998) sowie Bearden und Teel (1983a) sprechen dafür, dass Kunden, die sich beschwerten, mit hoher Wahrscheinlichkeit wieder beim selben Anbieter kaufen werden, selbst wenn dessen Wiedergutmachung für einen Dienstleistungszwischenfall mangelhaft war: *“Complaining customers are among the most loyal”* (Eccles und Durand, 1998, S. 68).

Gemäss Bearden und Teel (1983b) wird Kundenbeschwerdeverhalten ausgelöst durch Emotionen der Unzufriedenheit. Aus soziodemografischer Sicht neigen ältere Konsumenten eher zur Beschwerdeeinreichung, da sie über mehr Freizeit verfügen und zudem glauben, dass die Firma sich für den Fehler erkenntlich zeigen wird (Solomon, 2006; Speer, 1996).

Die grösste Barriere für eine effiziente Wiederherstellung der Beziehung zum Kunden ist, dass nur fünf bis zehn Prozent aller unzufriedenen Kunden sich bei einem Zwischenfall beschwerten (Kotler, 2006; Chung und Hoffman, 1998; Tax und Brown, 1998). Stattdessen wechseln die meisten enttäuschten Kunden zur Konkurrenz und / oder erzählen Bekannten von der schlechten Erfahrung (Singh, 1990a).

Kunden, die sich beschwerten, erwarten in der Regel Kompensationszahlungen: *“Negative feedback is likely to be perceived (by the customer) as offering higher potential for compensation”* (Söderlund, 1998, S. 175). Kunden erwarten nach der Einreichung einer Beschwerde, dass die Firma Verantwortung für das Problem übernimmt und dieses löst. Die Fairnesswahrnehmung nach der Wiedergutmachung ist wichtiger als die Diskonfirmation von Erwartungen der Wiedergutmachung (Andreassen, 1999).

2.2 Strafreorien im Rechtssystem

Die Rechtslehre setzt sich intensiv mit dem Thema der Bestrafung auseinander. Sie liefert im Rahmen dieser Arbeit Einsichten zur Begriffsabgrenzung, zur Notwendigkeit, zum Sinn und zur Funktion von Strafen in Zivilgesellschaften. Da das Wirkungsmodell in dieser Arbeit auf die individuelle Wahrnehmung und Reaktion von „bestraften“ Konsumenten ausgerichtet ist, trägt die juristische Betrachtung nicht direkt zur Erarbeitung des Modells bei.

Das klassische Feld der Bestrafung ist das materielle² Strafrecht als öffentliches Recht. In der Schweiz ist es geregelt im Strafgesetzbuch (StGB) und einschlägigen Nebengesetzen. Daneben gibt es Strafen im privatrechtlichen Bereich. Um diese beiden Erscheinungsformen der Strafe voneinander anzugrenzen, sind zunächst die Grundlagen für die Anwendung von Strafen im eigentlichen Strafrecht erörtert (Kap. 2.2.1 - 2.2.3). Der privatrechtliche Bereich wird in Kapitel 2.2.4 erläutert. Da in jedem Land ein anderes Rechtssystem herrscht und die Analyse verschiedener Rechtssysteme den Rahmen dieser Arbeit sprengen würde, wird in diesem Kapitel basierend auf dem Schweizer Strafrecht argumentiert. Ein kurzer Vergleich mit deutschem und amerikanischem Strafrecht ist im Anhang dargestellt (A.1).

2.2.1 Notwendigkeit, Sinn und Funktion des Strafrechts

Der Wunsch nach Gerechtigkeit gehört zu den Grundbedürfnissen des Menschen, neben demjenigen der Freiheit (Kunz, 2003). Soziales menschliches Leben in Freiheit ist aber nicht denkbar ohne Regeln und Normen, deren Missachtung Sanktionen auslösen. Das Strafrecht dient in diesem Sinne dem Schutz sogenannter Rechtsgüter, d.h. individuelle oder allgemeine Werte, die für das friedliche menschliche Zusammenleben bedeutsam sind (Rehberg und Donatsch, 2001).

Bereits im 19. Jahrhundert sprach Emile Durkheim von Strafen im Zusammenhang mit der öffentlichen Meinung: „Die öffentliche Meinung verhindert jeden Akt, der sie verletzt [...] durch die besonderen Strafen, über die sie verfügt“ (Durkheim, 1895, S. 106). Für Durkheim sollte die Intensität der strafrechtlichen Reaktion ausschliesslich nach der Intensität des im Verbrechen verletzten Gefühls richten. Somit stehen nicht Funktionen wie *Abschreckung*, *Rache* oder *Rechtsgüterschutz* im Vordergrund, sondern die latente Funktion von Strafen, die *emotiven* Grundlagen sozialer Organisationen aufrechtzuerhalten (Gephart, 1990). Im Vergleich zur heutigen Rechtsprechung argumentierte der Soziologe Ferdinand Tönnies (1887) zeitgemässer als Durkheim. Er forderte, dass die Bestrafung nach der *Schwere* einer Tat gerichtet sein sollte, wobei unabhängige Richter nach *objektiven Massstäben* beurteilen sollen.

Es gibt in unserer sozialen Ordnung verschiedene Möglichkeiten, den Verstoss gegen Verhaltensregeln zu sanktionieren: z.B. Zwang zur Herstellung oder Wiederherstellung eines rechtmässigen

² Materielles Strafrecht befasst sich mit der Frage, ob das Verhalten von Personen strafbar ist und welche Sanktionen zu verhängen sind (Riklin, 2002). Dies steht im Gegensatz zum formellen Strafrecht, das sich beschränkt auf die gerichtliche Feststellung und Durchsetzung des materiellen Rechts.

Zustandes, Auflösung von Rechtsverhältnissen oder Verpflichtung zu Schadenersatz oder Genugtuung (Kunz, 2003). Aus Sicht des öffentlichen Strafrechts erscheint die Strafe als „eine besondere Form staatlichen Zwanges, die in erster Linie bezweckt, dem Betroffenen ein Übel zuzufügen, ihm eine mehr oder weniger drastische Einbusse an Lebensgütern aufzuerlegen“ (Stratenwerth, 2005, S. 3). Die Strafe ist daher ein empfindlicher Eingriff in die Rechtsgüter der normwidrig handelnden Person und die schärfste Form des Ersatzzwanges³. Sie ist *ultima ratio* und für schwere Rechtsverletzungen vorzubehalten (Rehberg und Donatsch, 2001). Letzteres lässt sich aber nicht immer verwirklichen, weil teilweise für geringfügige Normverstösse keine adäquate Form des Ersatzzwanges möglich ist (z.B. Strassenverkehr).

Einverständnis herrscht darüber, dass die Strafe nur noch rational mit dem Zweck des Schutzes der rechtlichen Ordnung begründet werden kann (Seelmann, 2005). Nach wie vor umstritten ist dagegen, welches Mass an Strafe dieser Zweck erfordert und auf welche Weise die Strafe ihn erfüllen mag (Stratenwerth, 2005).

2.2.2 Voraussetzungen für Bestrafung gemäss materiellem Strafrecht

Der Strafanspruch im öffentlichen Recht steht ausschliesslich dem Staat zu. Bestrafung setzt deshalb zwei zentrale Grundvoraussetzungen voraus: das Legalitäts- und das Schuldprinzip.

Strafbar ist nur, wer eine Tat begeht, die das Gesetz ausdrücklich mit Strafe bedroht (keine Strafe ohne Gesetz - *nulla poena sine lege*). Dies ist das Legalitätsprinzip, das praktisch in allen Kulturstaaten der Erde als Freiheitsgarantie durchgesetzt ist (Rehberg, 2001). Verbotenes Verhalten kann demnach nachträglich mit Sanktionen belegt werden, unerlaubtes Verhalten kann nicht durch nachträgliche Sanktionen zu verbotenem gemacht werden.

Die zweite Grundvoraussetzung ist das Schuldprinzip: Keine Strafe ohne Schuld (Art. 10 StGB). Strafrechtliche Haftung ist stets Haftung für verschuldetes Unrecht. Ist ein Widerspruch zur strafrechtlichen Sollensordnung erstellt, geht es auf der Wertungsstufe der Schuld um die Frage, ob der Beschuldigte für sein Zurückbleiben hinter der Sollensordnung nach strafrechtlichen Massstäben verantwortlich gemacht werden kann (Honsell et al., 2003).

2.2.3 Funktion der Strafe

Strafe als Vergeltung

Relative Straftheorien gehen von Erfordernissen der Prävention aus. Sozialwissenschaften können heute nachweisen, dass der Einzelne in einem bestimmten Zeitpunkt niemals beliebig entscheiden kann. Dies, weil sich menschliches Verhalten aus einem komplexen Zusammenspiel individueller

³ *Ersatzzwang* sind Vorkehrungen, die nachträglich versuchen, den durch die Norm angestrebten rechtmässigen Zustand wiederherzustellen: Konfiszierung eines fahruntüchtigen Autos, Eintreiben von nicht bezahlten Steuern, Schadenersatz für Zerstörung fremder Sachen etc.

und sozialer Faktoren ergibt (Carlsmith et al., 2002). Schuld im strafrechtlichen Sinne bedeutet daher nicht wirkliche Schuld und also auch keine Legitimation für eine schuldbezogene Strafe (Rehberg, 2001).

Die *absolute Straftheorie* hingegen geht von einer strengen Bindung der Strafe an die Schuld aus: das Mass der Strafe hat dem Verschulden des Täters zu entsprechen. Strafe wird als *Ausgleich der Tatschuld* oder *Ausdruck eines ethischen Unwerturteils* gedeutet (vgl. Art. 47 des StGB; Seelmann, 2005). Die absolute Straftheorie bezieht sich auf die äussere Vergeltung. Die Erregung der Öffentlichkeit besonders bei schweren Rechtsbrüchen enthüllt massive und gerechtfertigte Vergeltungsansprüche (Weiner et al., 1997). Die öffentliche Strafe erfüllt insoweit die Funktion, die Selbsthilfe des Einzelnen oder der sozialen Gruppe, das Faustrecht und die Lynchjustiz, abzulösen (Riklin, 2002).

Das zentrale Ziel von Bestrafung ist aus Sicht der Vergeltungstheorie, einen Täter des falschen Vorteils, den er sich durch die Verletzung von Normen schaffen wollte, zu berauben (Bradley, 2003). Bestrafung versichert einer Gesellschaft, dass sich kriminelles Handeln nicht lohnt und dass die Befolgung von Gesetzen wichtig ist. In diesem Sinne wird *Fairness* und *Equality* wieder hergestellt (Hart, 1982; vgl. Kap. 2.3.2). Vergeltung ist gemäss Bradley (2003) die einzige philosophische Basis für Bestrafung, weil sie auch Elemente der Abschreckung, Wiedereingliederung und Unfähigmachung beinhaltet. Gleichzeitig wird sichergestellt, dass die Unschuldigen geschützt werden und soziales Gleichgewicht wiederhergestellt ist, nachdem es gestört wurde durch eine Tat: *“Retribution is the only appropriate moral justification for punishment”* (Bradley, 2003, S. 31).

Strafe als Abschreckung

Der Strafe kann die Funktion zugeordnet werden, das Verbrechen durch Einwirkung auf den Täter selbst zu verhüten. Dies wird im Strafrecht Spezialprävention genannt. Entweder werden die kriminellen Verhaltensdispositionen des Täters beeinflusst (Abschreckung). Oder Täter werden durch äusseren Zwang (heute in Form von Internierung) an weiteren Delikten gehindert (Rehberg, 2001). Die Psychologie spricht in diesem Fall von utilitaristischer Bestrafung (vgl. Kap. 2.3.2). Sie ist eine Bestrafung mit dem Ziel, die Wahrscheinlichkeit eines Wiederholungsfalles zu reduzieren (Carlsmith et al., 2002).

Die Sanktionsforschung steht noch in den Anfängen. Unklar ist, ob die Strafe zur Rückfallverhütung notwendig oder auch nur tauglich ist (Stratenwerth, 2005). Untersuchungen weisen bspw. darauf hin, dass sich der Freiheitsentzug umso negativer auswirkt, je repressiver er ist (Kaiser und Schöch, 2003). Wenn vollständig auf Spezialprävention gesetzt würde, würde dies heissen, dass bei Tötungsdelikten auf Sanktionen verzichtet wird. Diesen Verzicht würde die Öffentlichkeit nicht hinnehmen (Kunz, 2003). Bedürfnisse der Spezialprävention können die Institution einer öffentlichen Strafe also nicht begründen. Ähnlich wie beim Schuldgedanken kann die relative Straftheorie der Spezialprävention nicht der Ausgangspunkt des Strafrechts sein (Riklin, 2002). Heute wird lediglich bei Entscheidungen über Art und Mass der zu verhängenden Strafe, bei der Form des Vollzuges,

bei der Anordnung strafrechtlicher Massnahmen etc. auf die Individualität des Täters eingegangen (Stratenwerth und Kuhlen, 2004).

In der öffentlichen Meinung und in der Rechtssprechung spielt die Generalprävention, die abschreckende Wirkung von Strafen, eine erhebliche Rolle (Seelmann, 2005). Sie ist im Strafrecht in neuerer Zeit in den Vordergrund getreten, obwohl Erkenntnisse der Sozialwissenschaft ihr nur eine partielle Anwendbarkeit bescheinigen. Einigermassen gesichert ist, dass die Geltung einer Norm davon abhängt, ob ihre Übertretung strafrechtlich geahndet wird. Auf Art und Mass der Sanktion scheint es innerhalb eines weiten Spielraums nicht anzukommen (Riklin, 2002). Zudem verbieten gemäss Seelmann (2005) prinzipielle Gründe, die Strafe allein auf den Abschreckungsgedanken zu stützen: *“These goals [deterrence and rehabilitation] may contribute to a sound account of punishment - they may be secondary aims of punishment - but none can, on its own, morally justify punishment”* (Bradley, 2003, S. 20, und Hart, 1982).

Das Schuldprinzip (persönliche Verantwortung des Einzelnen) wird der Generalprävention vorgezogen. Nur dieses Prinzip aus dem Strafrecht kann verhindern, wie es Kant formuliert, dass der Straftäter „bloss als Mittel“ benutzt und unter die „Gegenstände des Sachenrechts gemengt wird“ (Kant, 1797).

Strafe als Täter-Opfer-Ausgleich

Seit den früheren 80er Jahren des vergangenen Jahrhunderts wird der Strafe eine neue Aufgabe zugewiesen, die sich nicht mehr mit dem Schema Vergeltung oder Abschreckung erklären lässt. Mit dem Täter-Opfer-Ausgleich ist die Wiedergutmachung des durch die Straftat angerichteten Schadens gemeint. Durch Leistungen des Täters, die über den Schadenersatz hinausgehen, wird der soziale Konflikt bewältigt (Seelmann, 2005). Als Täter-Opfer-Ausgleich gilt dabei ein zwischen den Beteiligten persönlich (mit oder ohne Vermittlung) stattfindender Prozess der Verständigung über solche Leistungen.

Ein Rechtsbruch kann eine ganze Reihe von Interessen berühren und bei der Aufarbeitung müssen verschiedenste Bedürfnisse beachtet werden. Welche Bedürfnisse dies sind, hängt vom Einzelfall ab (Stratenwerth, 2005). Bei einer Tötung im Affekt kann es, anders als bei Wirtschaftsdelikten oder Kundenbestrafungen, nicht um Abschreckung Dritter, bei einem Chemieunfall nicht um Rückfallverhütung (bspw. bei Serientäter) und beim Drogenhandel nicht um Täter-Opfer-Ausgleich gehen (bspw. bei Vermögensdelikten).

Welche Verhaltensweisen mit Strafe bedroht, also *kriminalisiert* werden, darüber entscheidet in relativ weiten Grenzen ein sozialer Auswahlprozess, dessen Ergebnisse nicht im Vorherein feststehen (Riklin, 2002). Die Funktion von Strafe ohne den Rückgriff auf die realen gesellschaftlichen Bedürfnisse lässt sich nicht klären. Überall bestehen Beurteilungs- und Ermessensspielräume, die sich bei der Strafbemessung orientieren an: persönlichen Merkmalen der Betroffenen (Schichtzugehörigkeit, Alter, Geschlecht), am Charakter der Straftat, dem Opferverhalten, an Aufklärungs- und

Beweisschwierigkeiten oder an gängigen Auffassungen über die Gefährlichkeit bestimmter Delikte usw. (Kley-Struller, 1993).

2.2.4 Privatrecht und Strafe

Neben der öffentlichen Strafe gibt es Sanktionen wie Unternehmensstrafen, die nicht von behördlichen Instanzen, sondern von Privat- oder juristischen Personen verhängt werden. Im Gegensatz zum Strafrecht im öffentlichen Recht sind nicht die Interessen der Öffentlichkeit von Belang, sondern die Interessen zwischen zwei Parteien. Strafen im Privatrecht sind also kein Ausfluss des staatlichen Strafanspruches. Von strafrechtlichen Sanktionen kann jedermann betroffen sein. Private „Strafen“ setzen aber ein spezielles Verhältnis zwischen dem „Bestrafer“ und dem „Bestraften“ voraus. Dieses spezielle Verhältnis zueinander kann z.B. bestehen in einem Vertrag den ein Kunde unterschreibt, oder in einer Vereins- oder Verbandsmitgliedschaft.

Sanktioniertes *Unrecht* ist eine *Widerhandlung* gegen z.B. vertragliche oder statutarische Pflichten (z.B. Richtlinien in den Tarifbedingungen einer Airline). Je nachdem, welcher Bereich angesprochen ist, kann der mit der Privat-„Strafe“ verfolgte Zweck verschieden sein. Beispiele:

- Betriebsjustiz auf der Grundlage eines Arbeitsvertrages und dazugehörigen Reglementen und den sich daraus ergebenden Pflichten des Arbeitnehmers;
- Konventionalstrafen bei Verletzung von Vertragspflichten (Nichterfüllung, Schlechterfüllung - das Obligationenrecht kommt zur Anwendung);
- Vereins- oder Verbandsstrafen bei Verstoss gegen statutarische Bestimmungen. Grundlage ist die Vereins- und / oder Verbandsmitgliedschaft. Hier spielen vor allem Kartelle eine wichtige Rolle (Riemer, 1990).

Auch Privat-„Strafen“ müssen auf einer gültigen rechtlichen Grundlage beruhen (z.B. Arbeitsvertrag und dazugehörige Betriebsreglemente, Vereins- oder Verbandsstatuten).

Im Gegensatz zum StGB ist das Ziel der Privat-„Strafe“ nicht die Wahrung allgemeiner öffentlicher Interessen (Schutz der Rechtsgüter), sondern die Wahrung von Partikularinteressen: z.B. Durchsetzung von statutarischen und vertraglichen Pflichten, Durchsetzung von Verhaltensanordnungen im Dienste eines geregelten Betriebslebens.

2.2.5 Zusammenfassung

Die Strafe wird aus juristischer Sicht unterteilt in Strafe aus öffentlichem Recht und privatem Recht (vgl. Tab. 2-4). Das Ziel der Strafe aus öffentlichem Recht lässt sich in Teilaspekten als *Vergeltung*, *Abschreckung* oder *Täter-Opfer-Ausgleich* darstellen. Im übergeordneten Sinne geht es um die Wahrung öffentlicher Interessen (Schutz der Rechtsgüter). Im Gegensatz dazu geht es bei der in dieser Arbeit untersuchten Privat-Strafe um Partikularinteressen, wie bspw. die Durchsetzung von vertraglichen Pflichten zwischen Unternehmen und Kunden.

Verschiedene Theorien aus dem öffentlichen Recht helfen, den Hintergrund von Bestrafungen aus juristischer Sicht zu betrachten: Strafe bleibt aus Sicht der absoluten Straftheorie schuldbezogene Vergeltung (Kunz, 2003). Ohne Schuld ist Strafe nicht zu rechtfertigen; sie ist die Voraussetzung, damit jemand für einen rechtlich missbilligten Geschehensablauf verantwortlich gemacht werden kann. Das Schuldprinzip vermittelt aber keinen Aufschluss über die einzelnen Sanktionen einer Strafe. Hier spielt die Spezialprävention mit rein: die Strafe sollte so geartet sein, dass sie weiteren Delikten desselben Täters möglichst entgegenwirkt oder sie zumindest nicht fördert (Riklin, 2002).

Vereinfacht gesagt wird bei Delikten zuerst festgelegt, was unrecht ist, dann wie darauf reagiert wird, der konkrete Fall und die Hintergründe werden beachtet und zuletzt versucht, eine Besserung zu erreichen.

Tabelle 2-4: Privatrechtliche und öffentliche Strafen im Schweizer Strafrecht

	Strafen im öffentlichen Recht	Strafen im privaten Recht
Rechtsquelle	Strafgesetzbuch (StGB) und strafrechtliche Nebengesetze	Zivilgesetzbuch, Obligationenrecht und einschlägige Nebengesetze
Instanz (Strafanspruch)	Behördliche Instanzen	Privat- oder juristische Personen
Interessen	Interessen der Öffentlichkeit (Schutz der Rechtsgüter)	Interessen zwischen zwei Parteien (Partikularinteressen)
Voraussetzung	Legalitäts- und Schuldprinzip	Spezielles Verhältnis zwischen „Bestrafer“ und „Bestrafter“ (z.B. Vertrag)
Strafzweck	Vergeltung, Abschreckung (General- und Spezialprävention), Täter-Opfer-Ausgleich	Einhalten der Vertrags- oder Wettbewerbsbestimmungen (Schutz oder Abschreckung), verursachte Kosten entgelten

Quelle: Eigene Darstellung in Anlehnung an Stratenwerth, 2005

2.3 Wahrnehmungs- und Entscheidungsprozesse von Konsumenten

Konzepte und Einsichten aus der Psychologie erlauben die Analyse des menschlichen Verhaltens (Laucken et al., 1996). Nach wie vor beruhen die Wirtschaftswissenschaften auf dem Modell des *homo oeconomicus* als rational und eigennützig handelnder Mensch. Moderne wirtschaftswissenschaftliche Forschung, die versucht, menschliches Verhalten zu erklären, baut aber zunehmend Erkenntnisse der Sozialpsychologie in die ökonomischen Modelle ein (Fehr, 2005).

Ziel des einleitenden Kapitels 2.3.1 ist, Einblick in die Prozesse der (Fairness-) Wahrnehmung zu geben. Der Zusammenhang und die Unterschiede zwischen Wahrnehmung, Kognition, Emotion, Motivation und Verhalten werden geklärt. Dabei wird Rückgriff genommen auf das Basismodell der Wahrnehmung und der Einstellungen. Dieses Modell wird gemäss vorliegenden Forschungsfragen angepasst und bildet die Grundlage zur empirischen Untersuchung verschiedener Kundensteuerungsmassnahmen. Zur Klärung der Wirkung von Strafen auf Individuen schliesst das Kapitel 2.3.1 mit der Lerntheorie der instrumentellen Konditionierung ab. Die in der vorliegenden Arbeit empirisch untersuchte Fairness-Bewertung spielt bei der individuellen Bewertung verschiedener Einflussfaktoren (Kundensteuerungsmassnahmen) eine zentrale Rolle. In Kapitel 2.3.2 werden die kognitiven Beurteilungen von Fairness basierend auf der *Justice Theory* und der darauf aufbauenden *Fairness Theory* bezüglich Bestrafungsereignissen aufgezeigt. Dabei spielt die *Attributionspsychologie* (Kap. 2.3.3) eine wichtige Rolle, denn sie erklärt, wie Menschen subjektive Meinungen bilden. Mit der *Prospect Theory* (Kap. 2.3.4) und der daraus abgeleiteten Verlustaversion wird das Entscheidungsverhalten von bestraften Individuen aufgezeigt. Erkenntnisse aus der *Prospect Theory* und der Attributionspsychologie werden zur Erarbeitung der Implikationen verwendet (vgl. Kap. 7.1).

Die psychologische Betrachtung schliesst mit Kapitel 2.3.5, den Folgerungen im Hinblick auf das provisorische Wirkungsmodell zu Wahrnehmungs- und Entscheidungsprozessen von Individuen.

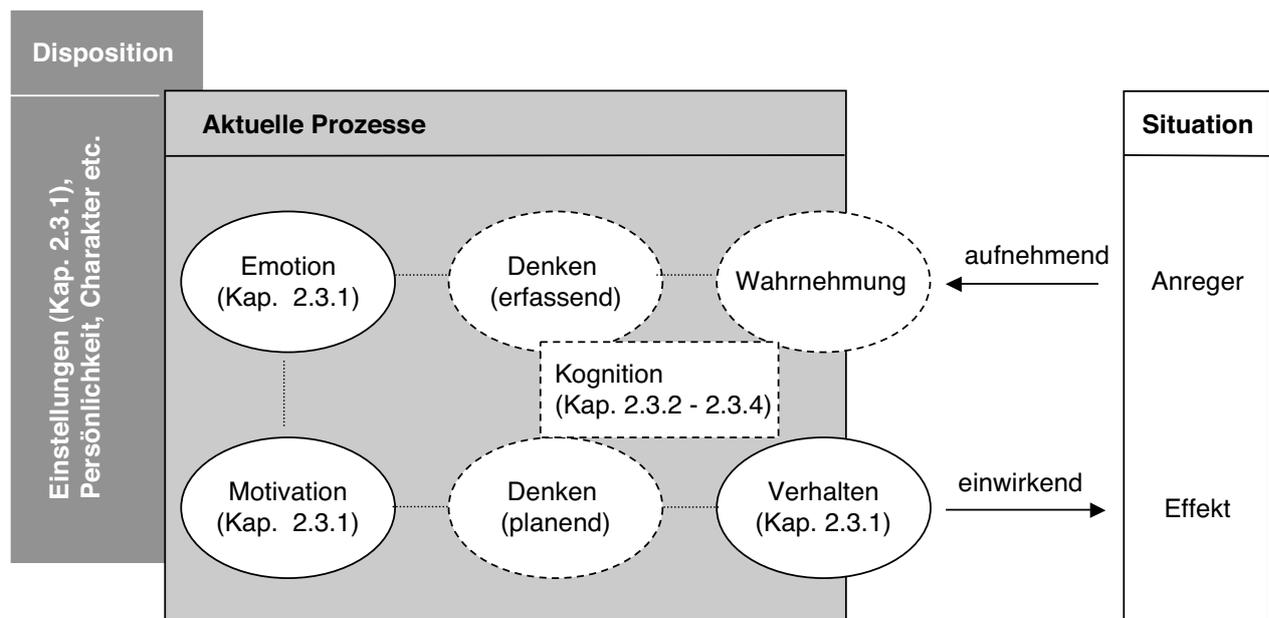
2.3.1 Basismodell zur Wahrnehmung

Um Wahrnehmungsprozesse und Verhalten besser verstehen zu können, wird in dieser Arbeit auf das Basismodell von Nolting und Paulus (2005) zurückgegriffen. Es weist Parallelen zum neobehavioristischen S-O-R-Verhaltensmodell (S-O-R: Stimulus-Organismus-Reaktion) auf, das in Modellen zum Käuferverhalten verwendet wird (Kuss und Tomczak, 2004). Der Behaviorismus (*behavior* = Verhalten) ist ein wissenschaftstheoretischer Standpunkt, bei dem das Verhalten von Menschen und Tieren mit den Methoden der Naturwissenschaften untersucht wird. Neobehaviorismus stützt sich zwar wie der klassische Behaviorismus auf Reiz-Reaktions-Beziehungen, enthält aber die weiter entwickelte Theorie über Reiz-Reaktions-Ketten, die durch klassische Konditionierung entstehen (Mills, 2000; O'Donohue, 1998).

Das Basismodell zur Wahrnehmung trägt zum Verständnis von Kognition und Emotion besser als das (reine) S-O-R-Modell bei, weil die inneren (aktuellen) Prozesse deutlich dargestellt werden. Es

ist auf einer Makrostruktur von *Aktuelle Prozesse*, *Personale Dispositionen* und *Situation* aufgebaut. Die Darstellung in Abbildung 2-2 zeigt die Zweiteilung der aktuellen Prozesse in Emotion / Motivation und Kognition, die sich in der Struktur von diesem Kapitel 2.3 widerspiegelt. Die aktuellen Prozesse, die personale Disposition und die Strafe als „Situation“ werden in den folgenden Abschnitten erläutert.

Abbildung 2-2: Basismodell zur Wahrnehmung



Legende: Kap. = Kapitel in der vorliegenden Arbeit (Einordnung der Themen in das Modell);
Quelle: Nolting und Paulus, 2005

Aktuelle Prozesse

Gemäss der Theorie des symbolischen Interaktionismus macht *Wahrnehmung* aus physikalischen Reizen eine Sache mit Bedeutung (Blumer, 1969). Dazu braucht es aber Bedeutungswissen, um bspw. einen Schlüssel (Symbol) als Instrument zum Öffnen von Türen zu erkennen (Schöpflug und Schöpflug, 1997). Wahrnehmung, Denken und Gedächtnisprozesse werden als *Kognition* bezeichnet. Kognition sind die internen Vorstellungen, die sich ein Mensch konstruieren kann (im Sinne des Konstruktivismus, vgl. Kap. 1.5.1).

Emotionen spielen bei Entscheidungen eine wichtige Rolle (Loewenstein et al., 2001). Der Emotion gilt in dieser Arbeit deshalb besondere Aufmerksamkeit. Es sind Vorgänge bzw. Zustände, die durch eine vorangehende Situation ausgelöst werden, und sie sind daher reaktiv. Typisch für Emotionen ist, dass sie ein Geschehen sind, das Individuen „überkommt“, und dass sie einhergehen mit einer leichten bis starken Erregung (Bagozzi et al., 1999). Wenn man einem Geschehen *unbeteiligt* gegenübersteht, werden keine Emotionen geweckt (Ulrich und Mayring, 1992). Wenn Emotionen Handlungsintentionen betreffen oder wenn sie Handlungen auslösen, die nicht oder kaum kontrollierbar sind, dann handelt es sich um Affekte (Bagozzi et al., 1999). Im Gegensatz zur Emotion geht

es bei der *Motivation* um den Antrieb: Eine Ziellage resp. ein Soll-Wert wird angestrebt (Rheinberg, 1995). Emotion und Motivation sind keine separaten Prozesse; es handelt sich vielmehr um zwei Seiten eines Prozesses (Kuhl, 1987). Ein Entschluss aufgrund einer Motivation wird nicht immer sofort in Handeln umgesetzt, sondern häufig erst gespeichert. Bei dieser Speicherung handelt es sich um eine Intention (Heckhausen und Heckhausen, 2006). Zur Übersicht und Definition der verwendeten Begriffe im Basismodell der Wahrnehmung, vgl. Tabelle 2-5.

Tabelle 2-5: Übersicht zu Kognition, Emotion, Motivation und Einstellungen

	Kognition	Emotion	Motivation	Einstellungen
Definition	Alle internen Vorstellungen: Wahrnehmung, Denken und Gedächtnisprozesse.	Reaktive Vorgänge, die durch eine Situation ausgelöst werden. Befindlichkeitsseite, die bei Entscheidungen eine wichtige Rolle spielt.	Antriebs-, Zielseite. Spannung zwischen momentaner und angestrebter Befindlichkeit. „Speicherung“ der Motivation ist eine „Intention / Absicht“.	Schemata, in denen Wissen und Gefühle gegenüber Objekten abgespeichert werden.
Anwendung in Arbeit	Theorien zur Erklärung kognitiver Vorgänge bei Bestrafungssituationen (für Wirkungsmodell).	Emotionen werden als „affektive Reaktion“ verstanden (Handlungsintention). Als abhängige Variable untersucht.	Intentionen sagen Kundenverhalten voraus: negative Verhaltensabsichten als abhängige Variablen untersucht.	Darstellung der Unterschiede und Zusammenhänge zwischen Kognition, Affekt und Verhalten.
Kapitel	2.3.2 – 2.3.4	2.3.1 und 2.5.3	2.1.4 und 2.3.1	2.3.1

Quelle: Eigene Darstellung

Anand et al. (1988) argumentieren mit dem *Cognitive-Affective Model*, dass die affektive Komponente resp. Beurteilung der letzte Schritt nach einer Serie von kognitiven Prozessen ist. Diese beinhalten die Sinneswahrnehmung von Stimuli und die Abfrage sinnvoller Information, um die Stimuli kategorisieren zu können (Anand et al., 1988). Eine andere Position nimmt Zajonc (1985) ein, der Affekt und Kognition als zwei unabhängige Systeme sieht. Affektive Reaktionen basieren nicht zwangsläufig auf kognitiven Prozessen. Diese Position findet Anwendung, wenn Kunden Produkte evaluieren und annehmen, dass diese angenehme Sinneswahrnehmungen auslösen (Solomon, 2006). Im zu erarbeitenden Wirkungsmodell dieser Arbeit wird der Argumentation von Anand et al. (1988) gefolgt, weil Wahrnehmung als der erste Schritt der inneren Grundprozesse angesehen wird. Dies wird ebenfalls im Basismodell dargestellt (vgl. Abb. 2-2). Auch Schoefer und Ennew (2005) gehen in ihrem Modell zur Fairness-Beurteilung nach einem Dienstleistungszwischenfall vom *Cognitive-Affective Model* aus.

Personale Disposition

Ein wichtiges Konzept bei der personalen Disposition sind die *Einstellungen* (vgl. Abb. 2-2). Sie sind Schemata, in denen das gesamte Wissen und die Gefühle gegenüber Einstellungsobjekten abgespeichert sind (Güttler, 2003). Nach dem Strukturmodell von Rosenberg und Hovland (1960)

kann jede soziale Einstellung in drei Komponenten aufgeteilt werden: *Kognitive*, *Affektive* und *Behaviorale Komponente*. Die drei Komponenten sind miteinander verbunden, voneinander abhängig und können als ein System bezüglich eines spezifischen Einstellungsobjektes aufgefasst werden (Güttler, 2003).

Um aus Einstellungen mögliches Verhalten (*Verhaltensabsichten*) vorauszusagen, wurden verschiedene Modelle und Theorien erarbeitet wie z.B. das *Fishbein Multiattribute Model* oder die *Theory of Reasoned Action* (Fischer und Wiswede, 2002). Im Bereich des Kundenverhaltens sind die Einstellungen eines Individuums jedoch keine guten Indikatoren, um künftiges Verhalten vorherzusagen. Einige Studien haben eine tiefe Korrelation zwischen Einstellungen und dem entsprechenden Verhalten nachgewiesen (vgl. Solomon, 2006).

Strafe als situative Anregung

Bei der Bestrafung, als eine Situation (Anreger) im Basismodell der Wahrnehmung, kann ein negativer Stimulus das Verhalten von Individuen verändern (Deutsch und Krauss, 1965). Um diesen Prozess besser verstehen zu können, eignet sich das von Skinner (1938) entwickelte Basiskonzept des instrumentellen Konditionierens (*Operant Conditioning*). Instrumentelles Konditionieren ist ein primärer Lernmechanismus im sozialen Verhalten. Es kann soziales Verhalten durch Einsatz von Stimuli fördern (Akers, 1977).

Während klassisches Konditionieren das Konzept des S-R-Lernens beinhaltet (Stimulus-Reaktion), geht es im instrumentellen Konditionieren um S-R-S (Stimulus-Reaktion-Stimulus)-Lernen. Das instrumentelle Konditionieren ist eine Weiterentwicklung des S-R Modells. Skinner fand heraus, dass die Wahrscheinlichkeit einer nochmaligen Reaktion (R) beim zweiten Stimulus (S) signifikant erhöht wird. Es reicht nicht als Stimulus, dem Kind nur eine Belohnung zu versprechen; das Versprechen sollte zur Verstärkung des Lerneffekts auch eingehalten werden.

Es gibt zwei Arten von Stimuli: positive und negative. Diese Stimuli können der Umwelt entweder zugefügt oder weggenommen werden, um die Wahrscheinlichkeit einer Reaktion (R) zu beeinflussen. Abbildung 2-3 zeigt das Verhältnis von Stimulus und Reaktion. Positive Verstärkung verstärkt Verhalten, indem ein positiver Stimulus hinzugefügt wird. Die negative Verstärkung führt zu einer Reaktion (R) durch Entfernung eines Stimulus. Die aversive Bestrafung Typ I führt zu einer Reaktion durch Hinzufügen eines negativen Stimulus und der Verlust (Bestrafung II) durch Entzug eines positiven Stimulus. Wenn das Verhalten ohne Konsequenzen bleibt, findet die *Löschung* statt.

Kunden, die durch Unternehmen Bestrafungen vom Typ I erlebt haben, werden ihr unerwünschtes Verhalten vermindern und sich den Unternehmensrichtlinien fügen (im Basismodell: Effekt). Bestrafungen des Typs II führen ebenfalls zur Vermeidung von bestimmtem Verhalten, weil z.B. Privilegien nach inadäquatem Verhalten weggenommen werden (Butterfield et al., 1996; van Houten, 1983). Das Konzept der positiven Verstärkung wird bewusst eingesetzt bei Loyalitäts- resp. Bonusprogrammen (Solomon, 2006). Wenn Kundenreaktionen von Unternehmen ignoriert werden,

Abbildung 2-3: Bestrafung durch instrumentelle Konditionierung

	... der Reizsituation neg. Stimuli	
	hinzufügen:	entfernen:
Positive Konsequenz ...	Positive Verstärkung	Bestrafung II
Negative Konsequenz ...	Bestrafung I	Negative Verstärkung

Quelle: Solomon, 2006

wird eine Löschung des Kundenverhaltens erwartet. Der Diskurs bricht ein und der Kunde wird sich je nach Situation in negative Verhaltensabsichten engagieren.

2.3.2 Fairness-Beurteilung

Die Fairness-Beurteilung bildet im Forschungsmodell dieser Arbeit die zentrale Wahrnehmungskomponente zur Beurteilung von Kundensteuerungsmassnahmen. In der empirischen Untersuchung wird die Fairness-Beurteilung einzelner Pull- oder Push-Massnahmen von den Probanden auf einer Rating-Skala eingestuft. Im Modell wird davon ausgegangen, dass eine Anreiz-Option (Pull-Massnahme) zu einer hohen Fairness-Beurteilung führt. Eine Richtlinie mit Bestrafung (Push-Massnahme) hingegen führt zu einer tiefen Fairness-Beurteilung. Der theoretische Bezugsrahmen zur Fairness-Beurteilung wird von der *Justice Theory* (Theorie der Gerechtigkeit) und der *Fairness Theory* (Fairness-Theorie) gebildet. Zur Übersicht zu den in der vorliegenden Arbeit verwendeten Elemente aus der *Justice* und *Fairness Theory*, vgl. Tabelle 2-6.

Die *Justice Theory* beschäftigt sich mit kognitiven Prozessen. Sie wurde in verschiedenen Bereichen angewandt, vor allem bei Untersuchungen zu Dienstleistungszwischenfällen, Kundenbeschwerden und monetären Bestrafungen (vgl. Wirtz und Mattila, 2004; McCarthy und Fram, 2000; Tax et al., 1998; Bitner et al., 1990; Oliver und Swan, 1989): “[...] *perceived justice / fairness is increasingly identified as a key influence in the formation of consumer’s evaluative judgements* [...]” (Schoefer und Ennew, 2005, S. 261). Verschiedene Autoren betonen die Wichtigkeit der Fairness-Beurteilung für die emotionale Reaktion, die Bewertung von Kundenzufriedenheit und für die Verhaltensabsichten (Kim, 2003; Maxham und Netemeyer, 2002; Smith et al., 1999; Weiss et al., 1999; Oliver und Swan, 1989).

Die *Justice Theory* besteht im Kern aus drei Dimensionen - die *distributive*, die *prozedurale* und die *interaktionale Gerechtigkeit*, welche sich gegenseitig beeinflussen (Cropanzano und Folger, 1991). Während in der Forschung zu Bestrafungen und kritischen Ereignissen meist von diesen drei Gerechtigkeitsskonzepten ausgegangen wird (z.B. Kim und Smith, 2005; Wirtz und Mattila, 2004),

beschäftigt sich die Sozialpsychologie mit einer vierten Dimension, der *restorativen Gerechtigkeit* als Wiedergutmachung (Gilovich et al., 2006). Im Dienstleistungsmarketing wird auch von *Service Recovery* gesprochen. Sie wird für den Airline-Bereich in Kapitel 3.1.7 abgegrenzt.

Equity und Equality (distributive Gerechtigkeit)

Das *Equity*-Prinzip basiert auf *Ratio Equivalence*, wonach Individuen Auszahlungen entsprechend ihrem Einsatz erhalten sollten. Es ist ein multidimensionales Konstrukt, das auch als „Fairness“ bezeichnet wird (Burns et al., 2000). Eine Strafzahlung wird aus der Equity-Sicht als unfair empfunden, wenn der Kunde den ihm zugefügten Schaden durch die Bestrafung als grösser wahrnimmt als der Schaden, den er der Unternehmung zugefügt hat (Kim, 2003). Der Einbezug des *Ticketpreises* als Einflussfaktor im vorliegenden Forschungsmodell berücksichtigt die Equity-Wahrnehmung. Es wird angenommen, dass das Verhältnis zwischen dem Ticketpreis und der Strafgebühr (oder Anreiz-Option) Equity-Überlegungen auslöst. Entscheidet sich der Kunde bspw., sich zu beschweren, findet eine Equity-Überlegung statt. Die potentielle Kompensation bei einer Beschwerde und die Kosten, um sich zu beschweren, sind analog zu den *Outcomes* und *Inputs* der Equity-Sicht. Beschwerden sind damit eine Reduktion der *Inequity* (Deutsch, 1975).

Wenn eine Person *Inequity* wahrnimmt, entsteht Unzufriedenheit oder andere negative emotionale Gefühle wie Ärger, Wut oder Schuld. Diese emotionalen Reaktionen lösen beim Individuum Handlungen aus, um die Equity wiederherzustellen (Lapidus und Pinkerton, 1995). Individuen ziehen ausgewogene Beziehungen gegenüber unharmonischen vor (Theorie der kognitiven Dissonanz; Festinger, 1957). Methoden, um die wahrgenommene *Inequity* zu neutralisieren, sind: Wechsel des Inputs in die Beziehung, Wechsel der Wahrnehmung des *Outcomes* oder Beenden des Verhältnisses (Walster et al., 1973).

Beim Prinzip *Equality* geht es um die gleichmässige Aufteilung einer Auszahlung auf die partizipierende Gruppe. Die Equality bezieht sich somit auf Drittpersonen. Die Frage lautet, ob die Bestrafung für alle Kunden mit demselben Tatbestand gleich ist (vgl. Kap. 2.4.3).

Prozesse (prozedurale Gerechtigkeit)

Drei Faktoren beeinflussen gemäss Tyler (1994) die Wahrnehmung der prozeduralen Gerechtigkeit. Erstens dürfen keine Bedenken über die *Neutralität* der Autorität bestehen. Zweitens muss *Vertrauen* zwischen dem Individuum und der Autorität bestehen. Individuen müssen das Gefühl haben, gemäss denselben Standards und Prinzipien behandelt zu werden. Drittens will der Mensch fair und mit *Respekt* behandelt werden.

Die Beurteilung von Neutralität, Vertrauen und Respekt sind gemäss Tyler (1994) stärkere Determinanten der Fairness-Wahrnehmung als die Belohnung oder Bestrafung selbst. Demnach werden bspw. Kündigungen oder monetäre Bestrafungen von Individuen ohne Protest akzeptiert, solange sie von einer Autorität in neutraler, vertrauenswürdiger und respektvoller Art erfolgen (Gilovich

et al., 2006). Die Geschwindigkeit, mit welcher Dienstleistungsfehler korrigiert oder Beschwerden behandelt werden, ist eine der Hauptdeterminanten von prozeduraler Gerechtigkeit (Tax et al., 1998). Prinzipien der prozeduralen Gerechtigkeit sind in Bestrafungssituationen anwendbar, weil Kunden die Prozesse evaluieren, wenn sie einer Bestrafungssituation ausgesetzt sind.

Tabelle 2-6: Kognitive Elemente zur Fairness-Beurteilung

	Inhalt	Beispiele bei Bestrafung
Distributive Gerechtigkeit	Verhältnis zwischen Auszahlungen und Einsatz (Equity). Aufteilung der Auszahlung (Equality).	<ul style="list-style-type: none"> – Ist Bestrafung fair im Verhältnis zum Kaufpreis? – Werden alle Kunden nach denselben Richtlinien bestraft?
Prozedurale Gerechtigkeit	Fairness des Verfahrens und der Richtlinien.	<ul style="list-style-type: none"> – Kann Bestrafung an die Situation angepasst werden (z.B. Kunde kann Situation nicht beeinflussen)?
Interaktionale Gerechtigkeit	Kommunikation der Strafe.	<ul style="list-style-type: none"> – Wird dem Kunden der Grund für die Bestrafung mitgeteilt? – Wurde korrekt, klar und anständig kommuniziert?
Restorative Gerechtigkeit	Entschuldigung (Versöhnung).	<ul style="list-style-type: none"> – Hat sich die Firma für das Verhalten entschuldigt?
Kontrafaktisches Denken	Interne Vorstellung zu Konsequenzen eines anderen möglichen Ereignisses.	<ul style="list-style-type: none"> – Gibt es verschiedene Arten, wie die Firma auf die Bestrafungssituation reagieren könnte?
Verantwortlichkeit	Zuordnung der Verantwortlichkeit.	<ul style="list-style-type: none"> – Kann Firma für Verhalten verantwortlich gemacht werden?

Quelle: Eigene Darstellung in Anlehnung an Tax et al., 1998

Kommunikation (interaktionale Gerechtigkeit)

Bies und Shapiro (1987) definieren die interaktionale Gerechtigkeit mit „Gerechtigkeit innerhalb der interpersonalen Kommunikation im Austausch verschiedener Parteien“. Sie ist gemäss Folger (1993) genauso wichtig wie die distributive und prozedurale Gerechtigkeit. Dürftige interpersonale Behandlung von Kunden oder fehlende Entschuldigung vermindert jeglichen positiven Effekt, z.B. einer hohen Kompensationszahlung (Tax et al., 1998). Elemente der interaktionalen Gerechtigkeit sind: Erklärung über Grund der Strafe, Ehrlichkeit, Anstand, Einsatz und Empathie (Smith et al., 1999; Tax et al., 1998; Goodwin und Ross, 1992; Bies und Shapiro, 1987). Interaktionale Gerechtigkeit wird in Bestrafungssituationen angewendet, weil Evaluationen von Kunden davon abhängen, wie Mitarbeiter des Dienstleistungsunternehmens mit ihnen kommuniziert haben. Die *Kommunikation* einer Kundensteuerungsmassnahme wird als Einflussfaktor (Variable) im Forschungsmodell identifiziert (deutliche vs. undeutliche Kommunikation).

Kontrafaktisches Denken und Verantwortlichkeit

Die *Fairness Theory* integriert Ergebnisse der Gerechtigkeitsforschung (Folger und Cropanzano, 1998), des kontrafaktischen Denkens (*Contrafactual Thinking*) und der Verantwortlichkeit (*Accountability*). Sie geht davon aus, dass aufgrund kognitiver Prozesse bezüglich Verantwortlichkeit eine negative Wahrnehmung von Fairness entsteht (Folger und Cropanzano, 1998). Für jegliche Ungerechtigkeit spielt in der *Fairness Theory* die Wahrnehmung, dass jemand dafür schuldig resp. verantwortlich ist, eine zentrale Rolle. Für diese kognitive Leistung ist kontrafaktisches Denken nötig. Es ermöglicht Menschen die Vorstellung, was für Konsequenzen ein anderes mögliches Ergebnis nach einem negativen Ereignis gehabt hätte (Roese, 1997). Vor allem negative Ereignisse lösen kontrafaktisches Denken aus (Morris und Moore, 2000).

Die individuelle Beurteilung eines Zwischenfalls resultiert in einer kognitiven Evaluation der Verantwortlichkeit / Fairness und einer emotionalen Antwort (McCull-Kennedy und Sparks, 2003; Folger und Cropanzano, 1998). Wenn sich ein Kunde viele mögliche Handlungsoptionen des Dienstleistungsanbieters nach einem Dienstleistungszwischenfall vorstellen kann und keine davon gewählt wird, resultiert dies in Unzufriedenheit oder Wut beim Kunden (McCull-Kennedy und Sparks, 2003). Diese starken negativen Emotionen entstehen vor allem dadurch, dass Kunden in diesem Fall merken, wie wenig Respekt und Effort von Seiten des Dienstleistungsanbieters aufgewendet werden. Negative Emotionen spielen bei der *Fairness Theory* insofern eine wichtige Rolle, als dass sie sowohl kritisches, kontrafaktisches Denken auslösen, aber auch Resultat des Fairness-Beurteilungsprozesses sind (McCull-Kennedy und Sparks, 2003; Roese, 1997).

2.3.3 Attributionstheorie

Kritiker der Fairness-Theorie schlagen vor, Ergebnisse der *Attributionspsychologie* als Ergänzung zu nutzen, denn die Equity ist bspw. nur eine von vielen möglichen Normen hinter dem Gerechtigkeitskonstrukt (Utne und Kidd, 1980). Eine Attribution ist eine subjektive Meinung eines Menschen über eine Kausalbeziehung (Heider, 1958). Das Konzept der Attributionen wird im Dienstleistungsmarketing eingesetzt, um zu erklären, wie Menschen Dinge um sich herum interpretieren, vor allem bei Ereignissen, die sie als negativ wahrnehmen (Swanson und Kelley, 2001; Blodgett et al., 1993; Weiner, 1985). Die Attributionstheorie hilft dabei, Gründe für das Kundenverhalten und für Attributionsverzerrungen zu verstehen (vgl. Tab. 2-7; Weiner, 2000; Bitner et al., 1994; Peeters und Czapinski, 1990).

Weiner (2000) hat eine Kategorisierung von Attributionen entwickelt, welche die Gründe für ein Produkt- oder Dienstleistungsfehler nach drei Dimensionen klassifiziert: (a) *Lokus* der Kausalität, (b) *Stabilität* und (c) *Kontrollierbarkeit* (vgl. Tab. 2-7). Bei (a) handelt es sich um die Zuordnung der Attribution entweder auf sich selbst (z.B. Persönlichkeit) oder auf externe Faktoren (z.B. Wetter). Für den Bereich von Bestrafungen heisst dies, dass die Kundenzufriedenheit tiefer sein wird, wenn die Bestrafungen externen Faktoren zugeordnet werden können, als wenn sie internen Fakto-

Tabelle 2-7: Attributionen im Kontext von Strafen

Attribution		Attributionsverzerrungen
Lokus der Kausalität:	Zuordnung der Strafauslösung auf sich selbst oder auf externe Faktoren.	– Bedürfnis nach Kausalzusammenhängen; – Bei Erfolg wird internal attribuiert (self serving bias), Misserfolg externen Ursachen zugeordnet;
Stabilität:	Bewertung der Wiederholungswahrscheinlichkeit einer Ursache oder Strafe.	– Ereignisse werden verschieden wahrgenommen (ungleiche Attributionen);
Kontrollierbarkeit:	Anteil von Kontrolle durch Person oder Unternehmung an Auslösung der Strafe.	– Eigenes Handeln wird als normkonform betrachtet.

Quelle: Eigene Darstellung in Anlehnung an Weiner, 2000

ren zugeordnet werden. Kunden, die extern attribuieren, erzählen anderen eher über die schlechten Erfahrungen, als Kunden, die intern attribuieren (Richins, 1983). Die Dimension (b) betrifft die Dauerhaftigkeit resp. Wahrscheinlichkeit einer Wiederholung der Ursache. Im Falle von Bestrafungen handelt es sich um die Frage, ob in Zukunft nochmals mit einer Strafe oder einem kritischen Ereignis gerechnet werden muss (Folkes, 1984). Die letzte Dimension (c) betrifft den Anteil an Kontrolle, die eine Person oder eine Unternehmung an der Auslösung einer Strafe hat. Wenn negative Ereignisse entstehen, die ausserhalb der Kontrolle einer Unternehmung waren, so wird der Kunde nachsichtig reagieren. Er wird eher die Verantwortung übernehmen, wenn er bestraft wurde für eine Aktion, die er kontrollieren konnte (Taylor, 1994). Es spielt dabei keine Rolle, wie intensiv die Bestrafung wahrgenommen wird. Studien haben gezeigt, dass die Fairness-Wahrnehmung je nach Attribution von Lokus, Stabilität und Kontrollierbarkeit verschieden ist (Kim, 2003; Blodgett et al., 1993).

2.3.4 Prospect Theory

Da Bestrafungen oder negative kritische Ereignisse von Individuen als Verlust wahrgenommen werden und teilweise eine riskante Situation beinhalten (wenn versucht wird, der Strafe zu entgehen), eignet sich die *Prospect Theory* zur Erklärung von Entscheidungsverhalten in diesen Situationen (Kahneman und Tversky, 2000).

Entscheidungsprozess

Bei einem Entscheidungsprozess werden von Individuen die zur Auswahl stehenden Alternativen editiert. Die Erwartungen (*prospects*) werden so organisiert, dass die Entscheidungsfindung vereinfacht wird. Das Ergebnis eines Editierprozesses kann nicht im Vorhinein allgemein festgelegt werden. Dennoch haben Thaler et al. (1997) grundlegende Prinzipien aufgestellt, die sie *Hedonic Editing* nennen: Mehrere Gewinne sollten getrennt werden und mehrere Verluste sind zu kombi-

nieren; kleinere Gewinne sind von grösseren Verlusten zu trennen und kleinere Verluste sind in grössere Gewinne zu integrieren.

Aufgrund der *Prospect Theory* wird angenommen, dass eine schlechte Leistung bezüglich eines Attributes gute Leistungen anderer Attribute überwiegt (Mittal et al., 1998). Wenn bspw. ein Kunde unzufrieden ist wegen einer Bestrafung, so führt dies trotz guter Leistung anderer Attribute wie Freundlichkeit, Pünktlichkeit oder Verlässlichkeit zu einer schlechten Gesamtbeurteilung. Negativ gewertete Attribute (z.B. ungenügende Öffnungszeiten) haben einen grösseren Einfluss auf die Gesamtzufriedenheit, verglichen mit positiv gewerteten Attributen (Mittal et al., 1998).

Verlustaversion

Alternativen werden je nach Entscheidungsproblem in Bezug auf einen neutralen Referenzpunkt evaluiert (Kahneman und Tversky, 1979). Vom neutralen Referenzpunkt aus werden die möglichen Ergebnisse der Entscheidung als Gewinne beziehungsweise Verluste kodiert. Entscheidungen werden also nicht, wie von der traditionellen Ökonomie postuliert, mit Blick auf die Maximierung des Endvermögens, sondern unter Berücksichtigung der Veränderung relativ zum Status quo gefällt (Theil, 2002; Levy, 1996). Empirische Beobachtung von diesem Element sind *Framing-Effekte* von Entscheidungssituationen. Framing-Effekte verhindern gemäss Tversky und Kahneman (1981) die genaue Bestimmung des Referenzpunktes, resp. der entscheidungsrelevanten Grössen (z.B. wenn Fleisch anstatt „25% fett“ als „75% mager“ deklariert wird). Kahneman und Tversky liefern keine allgemeine Position zur Festlegung des Referenzpunktes. Sie zählen nur zwei Faktoren auf, die die Referenzgrösse im Alltag beeinflussen können: soziale Normen sowie Erwartungen. Für Unternehmen spielen demnach bei der Festlegung von Bestrafungsmechanismen der Kundensegmente die sozialen Normen und die Kommunikation der Produktattribute oder Richtlinien eine wichtige Rolle.

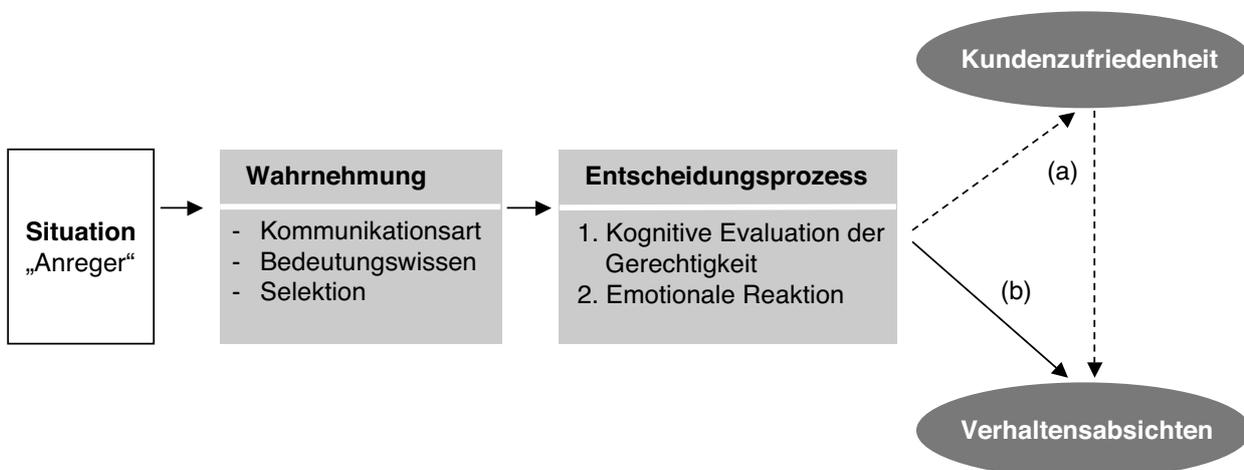
Güter im eigenen Vermögen werden höher bewertet als solche, die sich nicht in unserem Besitz befinden (Kahneman und Tversky, 1979): Auszahlungen - und folglich Strafgebühren - werden als Verluste begriffen. Die Haltung von Individuen gegenüber Verlusten ist stärker ausgeprägt als gegenüber Gewinnen. Verluste, im Besonderen hohe Verluste, werden gemäss empirischen Untersuchungen deutlich abgelehnt (Tversky und Kahneman, 1986). Das Prinzip der Verlustaversion findet nicht nur bei Entscheidungen unter Risiko Anwendung, sondern auch bei risikolosen Situationen (Tversky und Kahneman, 1991). Analysen des Ausmasses der Verlustaversion ergeben einen Faktor von ungefähr zwei zu eins: Verluste werden doppelt so stark wahrgenommen wie Gewinne (Kahneman, 1992).

2.3.5 Zusammenfassung

Aus sozial-psychologischer Sicht sind Strafen in der Beziehung zwischen Individuen gerechtfertigt, weil sie kooperatives Verhalten unterstützen und eine Möglichkeit geben, die Gerechtigkeit wiederherzustellen. Gemäss der Lerntheorie des instrumentellen Konditionierens gibt es verschiedene Bestrafungsarten, welche von Unternehmen eingesetzt werden können. Unternehmen müssen jedoch damit rechnen, dass die individuelle Fairness-Beurteilung dazu führt, dass eine als Strafe wahrgenommene Situation zu negativen Verhaltensabsichten führt.

Um eine Situation als Strafe oder Anreiz wahrzunehmen, wird im *Moment of Truth* Bedeutungswissen eingesetzt, während die Wahrnehmung als kognitive Leistung selektiv erfolgt (vgl. Abb. 2-4). Die Art der Kommunikation (korrekt, klar, anständig) nimmt dabei eine zentrale Rolle ein (*Framing*; Tversky und Kahneman, 1981). Der darauf folgende Entscheidungsprozess wird aufgrund von kognitiven Beurteilungen der Gerechtigkeit basierend auf den *Justice-*, *Fairness-* und *Attributionsprinzipien* durchgeführt. Die Situation wird gemäss der Fairness-Beurteilung evaluiert und eingestuft (Smith et al., 1999; Tax und Brown, 1998). In der Folge wird eine emotionale Reaktion hervorgerufen (Anand et al., 1988). Im Fall von negativen Ereignissen wie Bestrafungen wird die wahrscheinlichste Reaktion Unzufriedenheit, Frustration oder Wut sein (vgl. Kap. 2.5.2). Diese negativen Emotionen führen zu (a) Unzufriedenheit, Frust oder Wut und beeinflussen damit die negativen Verhaltensabsichten indirekt. Als affektive Reaktion führen sie (b) direkt zu Verhaltensabsichten (vgl. Abb. 2-4; Kap. 2.5.3). Wenn Kunden mit einer Bestrafung konfrontiert werden, führt dies zu Unzufriedenheit, auch wenn die anderen Attribute der Dienstleistung wunschgemäss erfüllt wurden (Mittal et al., 1998). Insbesondere negative Emotionen führen zu negativen Verhaltensabsichten (vgl. Kap. 2.1.4). Diese Erkenntnisse bilden die Grundlange zur Entwicklung des provisorischen Wirkungsmodells (vgl. Kap. 2.6).

Abbildung 2-4: Wahrnehmungs- und Entscheidungsprozesse



2.4 Strafen im Kontext zwischen Individuen und Organisationen

Ökonomische Theorien liefern Implikationen zur optimalen Gestaltung von Strafen aus der Anbieterperspektive. Unter den wenigen ökonomischen Theorien zur Bestrafung befindet sich die aus der Rechtssprechung abgeleitete Theorie der Abschreckung und Sättigung (Kap. 2.4.1). Durch Übertragung ökonomischer Logik auf soziale und gesellschaftliche Probleme tragen soziologische Theorien zum Erkenntnisgewinn bei (Swedberg, 1990). Die Organisationslehre (Kap. 2.4.2) gibt Antworten auf den Einfluss von Strafen auf die Zufriedenheit und Motivation von Mitarbeitern, während mit der Organisationssoziologie die Beziehung zwischen Individuen und Organisationen analysiert wird. Kapitel 2.4.3 behandelt die Gruppensoziologie, wobei die Frage im Vordergrund steht, wie das Individuum durch die Anwesenheit anderer beeinflusst wird. Die Erkenntnisse dieses Kapitels werden zur Bildung der Implikationen der vorliegenden Arbeit verwendet (vgl. Kap. 7.1).

2.4.1 Theorie der Abschreckung und Sättigung

Die ökonomische Theorie der Abschreckung wird in der Strafjustiz und in der Betriebswirtschaft eingesetzt: Die Aussicht auf eine Bestrafung für angerichtete Schäden hat eine abschreckende Wirkung auf negatives individuelles Verhalten (vgl. Kap. 2.2.3): „Wenn die Aktivität einer Person nicht wie erwartet belohnt oder unerwartet bestraft wird, wird die Person ärgerlich, und im Ärger sind die Ergebnisse aggressiven Verhaltens belohnend“ (Homans, 1972, S. 68).

Abschreckung und Anreize

Eine Strafe kann als eine Auferlegung von Kosten betrachtet werden. Ein Kunde wird somit negatives Verhalten unterlassen, wenn er weiss, dass er für dieses Verhalten (monetär) bestraft wird (Polinsky und Shavell, 1998; Kaplan et al., 1986). Menschen machen gemäss der Theorie der Abschreckung die Überlegung, ob sich der potentielle (monetäre) Erfolg einer Straftat lohnt, verglichen mit den Strafkosten und dem Risiko, erwischt zu werden. Becker (1968) folgert daraus, dass für den Staat die Maximierung der Strafhöhe die effizienteste Strategie ist, um Verbrechen zu verhindern. Wenn die Strafgebühren erhöht werden, entstehen nur marginale Kosten. Um die Wahrscheinlichkeit zu erhöhen, bei einer Straftat erwischt zu werden, entstehen hingegen überproportional hohe Kosten (z.B. erhöhter Einsatz der Polizei). Die optimale Bestrafung kann somit basierend auf ökonomischen Prinzipien erreicht werden (Sandmo, 1993).

Strafen zur Abschreckung werden vor allem im Bereich des Strassenverkehrs eingesetzt. Es hat sich gezeigt, dass nach der Einführung signifikant höherer Bussen die Geschwindigkeitsübertretungen abnahmen (o.V., 2000). Weil meistens Fehlverhalten oder eine Straftat nicht geahndet werden können, hat die Wahrscheinlichkeit, bestraft zu werden, einen grossen Einfluss auf das Konzept der Abschreckung. Die Doktrin der Abschreckung nimmt auch eine wichtige Rolle ein in der sozialen Kontrolle, wo auf Übereinstimmung mit Normen und Regeln geachtet wird (Meier und Johnson,

1977). Empirische Untersuchungen mit dem Ziel, einen Zusammenhang zwischen einer Strafpraxis und der Kriminalitätsentwicklung nachzuweisen, haben jedoch ergeben, dass der Einfluss von externem Druck noch unklar ist (vgl. Kap. 2.2.3). Die Anerkennung einer Richtlinie oder Norm hängt aber entscheidend davon ab, ob die Übertretung geahndet wird. Insofern hat das Konzept der Abschreckung in abgeschwächter Form seine Geltung behalten (Stratenwerth, 2005).

Einer Bestrafung zu entgehen, kann eine Belohnung (Anreiz) sein (Homans, 1972). Belohnungen erhöhen die Wahrscheinlichkeit, dass die erwünschte Handlung durchgeführt wird. Bestrafung dagegen erhöht die Wahrscheinlichkeit, dass eine Handlung, die eine Bestrafung verhindert, durchgeführt wird (Meleghy, 2001). Je „härter“ die in Aussicht gestellte Bestrafung, desto lohnender ist es, ihr zu entgehen. Was dabei als Belohnung und was als Strafe erlebt wird, ist zu einem grossen Teil im Prozess der Sozialisation gelernt (vgl. Meleghy, 2001).

Prinzip der Sättigung

Charakteristiken einer Interaktion zwischen Individuen können aus Sicht der *Social Exchange Theory* betrachtet werden (Homans, 1961). Basierend auf der *Theorie der instrumentellen Konditionierung* (vgl. 2.3.1) geht Homans davon aus, dass individuelles Verhalten durch Stimuli beeinflusst werden kann. Um Belohnung und Bestrafung zu bewerten, nutzen Individuen das Konzept der Sättigung, oder wie es Ökonomen nennen, *Prinzip des Marginalnutzens*. Je mehr Belohnung (Anreize) jemand erhalten hat, umso weniger wert wird jede zusätzliche Belohnungseinheit - die Nutzenfunktion ist konkav. Der marginale Nutzenwert einer Bestrafung läuft demgegenüber im negativen Bereich der Nutzenskala und ist ebenfalls konkav: Je mehr jemand bereits bestraft wurde, umso unangenehmer ist jede zusätzliche Bestrafungseinheit (Kim, 2003). In Übereinstimmung mit diesen Resultaten konnten Mittal et al. (1998) nachweisen, dass ein zusätzlicher Input einer positiven Leistung im Sinne des abnehmenden Grenznutzens weniger zur Gesamtzufriedenheit beiträgt. Wenn hingegen eine zusätzliche Einheit einer negativen Leistung vom Kunden wahrgenommen wird, dann trägt dieser Input zusätzlich zur negativen Nutzenwahrnehmung bei. Daher sollte gemäss Homans Theorie ein adäquater Level an Bestrafung nicht überschritten werden, weil sonst Kunden die Bestrafung als zu unfair beurteilen (Homans, 1961). Die Konsequenzen von *over-punishing* sind schädlich für Unternehmen, weil Kunden sich in negative Verhaltensabsichten engagieren. Diese negativen Verhaltensabsichten haben einen grösseren negativen Einfluss auf das Unternehmensergebnis als das Einkommen aus bspw. monetärer Bestrafung (Kim, 2003).

2.4.2 Individuen und Organisationen

In der *betriebswirtschaftlichen Organisationslehre* wird die Rolle von Bestrafung im Kontext Vorgesetzte-Mitarbeiter analysiert. Bestrafung wird hier definiert als ein „Verfahren der Organisation, bei der durch den Vorgesetzten negative Konsequenzen angewendet (oder positive Konsequenzen beseitigt) werden, um unerwünschtes Verhalten zu vermindern“ (Butterfield et al., 1996).

Mitarbeiter, die das Gefühl haben, sie hätten eine hohe Kontrolle über die Bestrafungsprozesse und die daraus resultierenden Folgen, engagieren sich in hohem Masse für die Organisation (Ball et al., 1994). Wenn Bestrafungen aber als zu hart wahrgenommen werden, hat dies einen negativen Einfluss auf die Leistung der Mitarbeiter. Die Bestrafungshärte sollte daher der Normverletzung angepasst werden. Wichtig für das Verhalten von Mitarbeitern nach einer Bestrafung ist die individuelle Fairness-Wahrnehmung (Ball et al., 1994).

Dass Bestrafungen sogar positive Effekte auf die Zufriedenheit, Motivation und Fairness-Wahrnehmung in Organisationen haben, konnten O'Reilly und Puffer (1989) nachweisen (vgl. Kap. 2.3.2). In gewissen *Low-Performance*-Kontexten hat das Fehlen einer Strafandrohung zur tieferen Motivation, schlechteren Equity-Wahrnehmung und Zufriedenheitsbewertungen von Mitarbeitern geführt. Richtig organisierte Bestrafung ist besser als keine Bestrafung, um Motivation, Zufriedenheit und Gerechtigkeitswahrnehmung zu erhöhen. Verglichen mit Strafandrohungen für schlechte Leistungen sind Anreize mit Belohnungen jedoch effizienter zur Erhöhung von Motivation und Zufriedenheit von Mitarbeitern (O'Reilly und Puffer, 1989). Dies ist konsistent mit Aussagen vom instrumentellen Konditionieren (vgl. Kap. 2.3.1), wonach positive Bekräftigung effektiver in der Beeinflussung von Verhalten ist als Bestrafung (Butterfield et al., 1996).

Die Resultate der Organisationsforschung sind nicht in jeder Situation anwendbar auf das zu erarbeitende Wirkungsmodell dieser Arbeit. In der Verhaltensforschung bei Organisationen wird auf Theorien des sozialen Lernens aufgebaut, wo es um den Lerneffekt bei der Beobachtung von Bestrafung bei anderen geht. In einer typischen Situation zwischen Unternehmen und Kunde kann aber gelerntes Verhalten nicht angewendet werden, weil andere Bestrafungen in der Vergangenheit nicht beobachtbar waren. Dennoch tragen Lerntheorien zum Verständnis des Kundenverhaltens bei, weil nach einem negativen Input sich das Verhalten des Kunden in Zukunft ändert, auch wenn die Bestrafungssituation einmalig ist.

Die *Organisationssoziologie* von Coleman (1990) liefert Antworten zur Frage, warum zwischen Individuen und Organisationen eine asymmetrische, unfaire Beziehung herrscht. In modernen Gesellschaften gibt es *individuelle* und *korporative Akteure*. Die Gesellschaft besteht aus Personen und Organisationen, wobei Letztere das gesellschaftliche Leben immer mehr beeinflussen. Coleman geht davon aus, dass individuelle Akteure kognitive Schwierigkeiten im Umgang mit korporativen Akteuren (z.B. Dienstleistungsunternehmen) haben. Im Zuge unserer Sozialisation wurden wir nicht auf Beziehungen mit korporativen Akteuren trainiert. Individuen wenden die Logik der erlernten Beziehungen zwischen Menschen auf die Beziehung mit Organisationen an (Coleman, 1982).

Zur Beurteilung der Fairness in einer Austauschbeziehung dient das Machtgefälle. Das Machtgefälle zwischen den beiden Akteurgruppen Individuum - Organisation liegt zu Ungunsten der individuellen Akteure (Parson und Sihls, 1962). Die Beziehung ist asymmetrisch, weil Unternehmen über viel mehr Ressourcen, Expertenwissen und mehr potentielle alternative Interaktionspartner verfü-

gen. Dies gibt den korporativen Akteuren die Möglichkeit, die Konditionen der Austauschbeziehung einseitig zu ihren Gunsten auszugestalten (Preisendörfer, 2005). Die systematische Benachteiligung der individuellen Akteure im Bereich der Möglichkeiten und Rechte auf Information ist für Coleman (1982) ein wichtiger Teilaspekt der Asymmetrie und Konfliktträchtigkeit zwischen Individuen und Organisationen.

2.4.3 Individuen in der Gruppe

Menschen verhalten sich in Gruppen anders (Fischer und Wiswede, 2002). Im Airline-Serviceprozess handelt das Individuum meist im Gruppenkontext und im Austausch mit Mitarbeitern einer Airline (vgl. Kap. 3.1.2). Einsichten aus der Gruppensoziologie tragen als theoretischer Rahmen dazu bei, den Einfluss anwesender Personen auf das individuelle Verhalten zu verstehen. In dieser Arbeit wird *Gruppe* in Anlehnung an Schäfers (1999) als die Gesamtheit der Personen verstanden, die sich ohne Verabredung am gleichen Ort aufhalten (z.B. die Warteschlange vor dem Check-In-Schalter).

Ursprünglich wurde angenommen, dass die Anwesenheit anderer das Erregungsniveau steigert und damit dominante Reaktionen hervorruft (Zajonc, 1965). Dies hilft beim Lösen von einfachen Aufgaben und hindert bei schwierigen Aufgaben (*triebtheoretische Erklärung*). Eine weitere Erklärung ist die *Bewertungsangst-Theorie*. Dabei steht die Sorge im Vordergrund, von anderen bewertet und mit deren Einstellungen, Urteilen konfrontiert zu werden (Girgensohn-Marchand, 1999). Die bloße Anwesenheit anderer genügt gemäss dieser kognitiven Theorie nicht. Die dritte Erklärung ist der *Aufmerksamkeitskonflikt*: ein Handlungshindernis entsteht durch mangelnde Konzentration, weil sich das Individuum zwischen dem Fokus auf die Aufgabe oder auf das Publikum entscheiden muss (Beckmann, 1992).

Merton (1957) geht von der Annahme aus, dass Menschen ihr Verhalten in einer Situation bestimmen, indem sie sich mit anderen vergleichen, sich abgrenzen und sich angleichen. Nicht selten taucht eine Überkonformität im beobachtbaren Handeln eines Individuums im Vergleich zur Bezugsgruppe auf (Fischer und Wiswede, 2002). Diese Überkonformität ist nicht nur mit dem Abbau der Dissonanz zu erklären (Irle, 1985; Festinger, 1957). Sie basiert im Wunsch, der Bezugsgruppe aufzuzeigen, dass die Normen und Werte gruppenkonform angewendet werden und auf richtigem *Impression Management* beruhen (Gukenbiehl, 1999). *Impression Management* ist der bewusste oder unbewusste, zielorientierte Versuch, die Wahrnehmung anderer Menschen während der sozialen Interaktion zu beeinflussen (Goffman, 1959). Aus der von Goffman (1959) entwickelten dramaturgischen Sicht ist menschliches Handeln eine permanente Inszenierung des einzelnen Akteurs. „Wir alle zielen vermittels unserer Selbst-Darstellung darauf ab, von den anderen auf eine bestimmte Art und Weise wahrgenommen zu werden, vor den anderen in einem bestimmten Licht zu erscheinen. Und die anderen machen im Prinzip dasselbe, und so machen wir alle uns sozial einander erträglich“ (Hitzler, 1992, S. 94).

2.4.4 Zusammenfassung

Die Theorie der Abschreckung hat gezeigt, dass hohe monetäre (oder „harte“) Bestrafungen helfen, unerwünschtes Kundenverhalten zu vermeiden (Becker, 1968). Dazu muss die Bestrafungshärte der Normverletzung angepasst werden, denn Kunden sind bis zu einem bestimmten Grenzbereich bezüglich Bestrafung indifferent. Wird jedoch die Grenze zum *over-punishing* überschritten, so wirkt sich das Strafmass schädlich auf die Unternehmung aus (Kim, 2003). Kundenreaktionen auf Strafen haben einen höheren negativen Einfluss auf das Unternehmensergebnis als der Nutzensgewinn des konformen Kundenverhaltens (Bougie et al., 2003). Auch die Wahrscheinlichkeit, bestraft zu werden, hat einen Einfluss auf die Wirkung der Strafe (Kaplan et al., 1986).

Konflikte zwischen Organisationen und Individuen basieren auf einem Machtgefälle (Coleman, 1982). Gerade im Falle von Bestrafungen besteht eine asymmetrische Beziehung. Individuen fühlen sich in der Austauschbeziehung einseitig schlecht behandelt und verfügen über zu wenig Informationen. Bestrafungen ziehen gemäss Organisationslehre nicht zwangsläufig negative Konsequenzen nach sich, wenn sie richtig umgesetzt werden. Sie tragen dazu bei, die Motivation und die Equity- sowie die Zufriedenheitswahrnehmung zu steigern. Trotzdem ist die positive Bekräftigung als „Anreiz mit Belohnung“ die effizientere Strategie zur Kundensteuerung (O'Reilly und Puffer, 1989). Die Fairness-Beurteilung und Aussagen der instrumentellen Konditionierung als Lerntheorie spielen dabei eine zentrale Rolle.

Zum Zeitpunkt eines negativen kritischen Ereignisses wird das Verhalten des Individuums vom Gruppenkontext beeinflusst. Anwesende Personen und die soziale Atmosphäre vor Ort nehmen Einfluss auf das Verhalten des Individuums (Girgensohn-Marchand, 1999).

Die Erkenntnisse aus diesem Kapitel werden zur Bildung der theoretischen Implikationen verwendet (vgl. Kap. 7.1).

2.5 Verwandte Konstrukte

Neben dem Bestrafungskonstrukt finden in der Forschung zum Kundenverhalten im Dienstleistungsbereich die Konstrukte *Diskonfirmation* (Kap. 2.5.1) und *Kundenzufriedenheit* (Kap. 2.5.2) Verwendung. Diese werden deshalb im Folgenden begrifflich abgegrenzt. Emotionen, als messbare Reaktion nach einer Kundensteuerungsmassnahme, werden anhand von Dienstleistungszwischenfällen in Kapitel 2.5.3 erläutert.

2.5.1 Erwartungs-Diskonfirmation

Erwartungs-Diskonfirmation ist eine kognitive Evaluierung, wobei der Kunde die Leistung eines Dienstleistungsunternehmens mit seiner Erwartungshaltung vor dem Konsum vergleicht (Kim und Smith, 2005). Gemäss dem *Expectancy Disconfirmation Model* formen Kunden Erwartungen basierend auf früherer Erfahrung und / oder Kommunikation durch die Unternehmung bezüglich der

Leistung eines Produktes oder einer Dienstleistung. Diese Überzeugungen implizieren einen gewissen Qualitätsstandard, der erwartet wird (Churchill und Surprenant, 1983).

Die Forschung sieht heute vom Erwartungs-Diskonfirmations-Paradigma ab (vgl. z.B. von Bartenwerffer, 2006; Spreng und Chiou, 2002; Wirtz und Mattila, 2001). Manche Forscher bezeichnen den Diskonfirmationsansatz sogar als „irrelevant“ oder „unnötig“ (Fournier und Mick, 1999; Cronin und Taylor, 1992). In dieser Arbeit wird davon ausgegangen, dass Kunden für Bestrafungen keine Erwartungen formen. Die Rolle von Diskonfirmation wird bei Bestrafungsereignissen geschwächt, denn *“customer(s) evaluate the objects about which they have not formed specific expectations”* (Kim und Smith, 2005, S. 175). Dienstleistungsunternehmen meiden eine effektive Kommunikation über *negative* Vorkommnisse wie Bestrafungen oder kritische Ereignisse. Daher haben Kunden bezüglich Bestrafungen keine spezifischen und klaren Erwartungen.

2.5.2 Kundenzufriedenheit

Viele Forscher postulieren ein direktes Verhältnis zwischen Zufriedenheit und anderen Konstrukten wie Loyalität und Dienstleistungsqualität oder differenzieren diese Konstrukte gar nicht (vgl. Athanassopoulos, 2000; Lee et al., 2000; Colgate und Stewart, 1998, etc.). Oliver (1997) hat vorgeschlagen, dass Kundenzufriedenheit das Resultat vom Vergleich zwischen Kundenerwartung mit dem Produkt / Dienstleistung sowie der tatsächlich gelieferten Leistung ist (vgl. Kap. 2.5.1). In einer späteren Arbeit hat Oliver (1999) Kundenzufriedenheit als *pleasurable fulfillment* definiert, das aus positiven Erlebnissen resultiert oder die Erwartungen übertrifft. Im Gegensatz zur Dienstleistungsqualität zeigt diese Definition, dass für die Evaluation von Kundenzufriedenheit mindestens eine Erfahrung mit der Dienstleistung nötig ist (vgl. Kap. 2.1.1). Es fehlt aber an Basiswissen und an einem gemeinsamen Verständnis sowie an Definitionen zum Zufriedenheitskonstrukt (White und Yu, 2005; Giese und Cote, 2000).

Einig sind sich verschiedene Forscher, dass die subjektive Evaluation von Emotionen zur Beurteilung von Kundenzufriedenheit zentral ist (vgl. White und Yu, 2005; Bagozzi et al., 1999; Wirtz und Bateson, 1999; Dubé, 1990). Bitner (1992) schlägt in diesem Zusammenhang vor, emotive Dispositionen als wichtige Mediatoren zwischen Dienstleistungsunternehmen und Kunden für die Verhaltensabsichten zu untersuchen. Wie die Theorien zur Fairness-Beurteilung gezeigt haben, können die emotionalen Reaktionen *Frustration* / *Wut* oder *Unzufriedenheit* nach der kognitiven Evaluation einer als unfair empfundenen Situation entstehen (vgl. Kap. 2.3.2). Wut und Frustration sind zuverlässige Indikatoren für Verhaltensabsichten (Stauss, 2004; Dubé und Menon, 2000).

In dieser Arbeit wird das Konstrukt *Kundenzufriedenheit (Customer Satisfaction)* als *ex post Beurteilung* eines Erlebnisses durch Kunden definiert. Es wird mit der Variablen *emotionale Reaktion* im provisorischen Wirkungsmodell abgebildet.

2.5.3 Emotionen bei der Dienstleistungsbegegnung

Kundenzufriedenheit reicht als Konstrukt nicht aus, um emotionale Komponenten und Reaktionen zu messen (Stauss und Neuhaus, 1997; Danaher und Haddrell, 1996). Bei der Dienstleistungsbegegnung spielen Emotionen von Konsumenten eine wichtige Rolle, weil der Kunde aktiv im Dienstleistungsprozess engagiert ist (Price et al., 1995). Vor allem bei einem negativen Ereignis wie einer Bestrafung wird beim Kunden eine negative emotionale Reaktion ausgelöst (vgl. Kap. 2.1.3).

Emotionen und Verhaltensabsichten

Die meisten Untersuchungen bestätigen eine hohe Korrelation zwischen der emotiven Evaluation und Verhaltensabsichten (vgl. Szymanski und Henard, 2001; Athanassopoulos, 2000; Zeithaml et al., 1996). Insbesondere wütende Kunden reagieren mit negativen Verhaltensabsichten: *“Events, particularly negative ones, are thought to produce emotions [...] Service Failure and recovery events have a significant impact on customer’s emotional and cognitive responses”* (McColl-Kennedy und Sparks, 2003, S. 254). Arbeiten im Bereich von Emotionen und Verhaltensabsichten haben gezeigt, dass positive Emotionen mit hoher Wiederkaufswahrscheinlichkeit und negative Emotionen mit tiefer Wiederkaufswahrscheinlichkeit assoziiert wurden (Babin und Babin, 2001; Machleit und Mantel, 2001; Andreassen, 2000). Eine starke negative Emotion führt gemäss den Untersuchungen von Bougie et al. (2003) und von Bagozzi et al. (1999) zu verschiedenen, gegen die Firma gerichteten Aktionen (vgl. Kap. 2.1.4):

- Umgehung (illoyales Verhalten, tiefe Wiederkaufswahrscheinlichkeit);
- Intensivierung des Efforts, um den inneren Konflikt zu entschärfen (negative WOM);
- Protest (Beschwerde einreichen).

Dabei muss aber unterschieden werden, ob ein Zwischenfall auf einem von der Firma kontrollierbaren Grund basiert oder nicht (vgl. Kap. 2.3.3). Kann diese Attribution durchgeführt werden und ist der Kunde wütend oder frustriert, so wird er sich mit hoher Wahrscheinlichkeit beschweren (Schneider und Bowen, 1999; Curren und Folkes, 1987).

Im vorliegenden provisorischen Wirkungsmodell wird die *emotionale Reaktion* als Variable aufgenommen, die einen direkten Einfluss auf die Verhaltensabsichten hat.

Messen von Emotionen

Da Emotionen schwierig zu kategorisieren sind, stellt sich die Frage nach der Messbarkeit. Emotionen beim Konsum von Dienstleistungen oder Produkten wurden bisher als eigenständige primäre Emotionen wie Freude, Interesse, Traurigkeit und Glück konzeptualisiert (Richins, 1997) oder als eine limitierte Anzahl von Basisdimensionen wie Vergnügen, Erregung oder positiver resp. negativer Affekt (Watson et al., 1988; Russell, 1980). Sie werden mit standardisierten Skalen wie der

Differentiated Emotion Scale (DES) (Oliver, 1993) oder mit der *Positive Affect Negative Affect Scale* (PANAS) (Watson et al., 1988) gemessen.

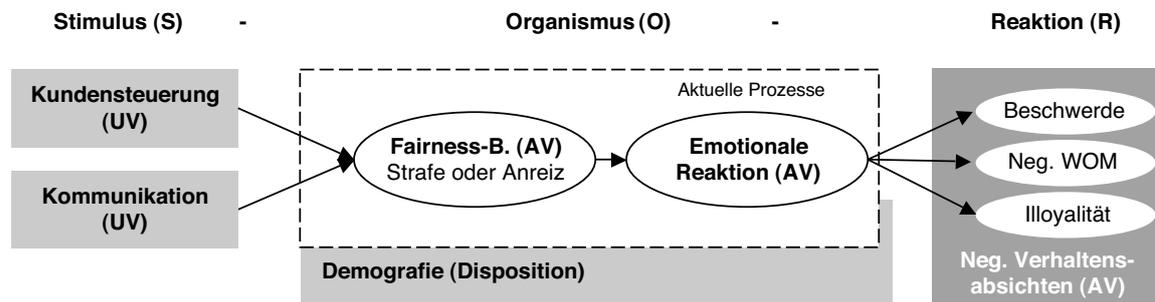
Mit Bezug auf Arbeiten von Holbrook und Batra (1987) sowie Edell und Burke (1987) schlagen Price et al. (1995) eine andere Skala vor. Sie unterteilen Emotionen in *positive* (glücklich, freudig, zufrieden, warmherzig, fürsorglich, herzlich, liebend) und *negative Emotionen* (traurig, betrübt, wütend, bedauernd). Diese Struktur hat sich in verschiedenen Konsumkontexten bewährt (Dubé und Menon, 2000). Sie wird für die vorliegende empirische Untersuchung verwendet und angepasst. Als Grundlage dient die Erkenntnis, dass Emotionen grundsätzlich in positive oder negative Emotionen kategorisiert werden können (White und Yu, 2005; Machleit und Mantel, 2001; Bagozzi et al., 1999; Liljander und Strandvik, 1997).

2.6 Provisorisches Wirkungsmodell

Aus *Anbieterperspektive* sind Kundenstrafen mit Richtlinien die Folge einer vom Kunden missachteten Einkaufsvereinbarung (McCarthy und Fram, 2000). Es sind Instrumente, um Kundenverhalten zu steuern oder zusätzlichen Umsatz zu generieren (Push-Massnahme). Dem Kunden werden aufgrund einer verpassten Opportunität negative Konsequenzen verursacht oder positive Konsequenzen verunmöglicht (Kumar et al., 1998). Als Alternative zur Kundensteuerung über Strafen eignen sich Anreiz-(Pull-) Massnahmen wie Bonusprogramme oder *Cross-* und *Upselling*. Mit der Steuerung über Anreiz-Massnahmen kann gleichzeitig der Kundenwert gesteigert werden (Lauer, 2004).

Aus *Kundenperspektive* bestimmt die Fairness-Beurteilung die kognitive Wahrnehmung und Einstufung der Kundensteuerungsmassnahme (Kim, 2006; Schoefer und Ennew, 2005; Maxham und Netemeyer, 2002). Das Resultat der Beurteilung kann im Falle einer Strafe zu einer starken Emotion wie Frust oder Wut führen, die als Aktionssignal negative Verhaltensabsichten auslöst (Bougie et al., 2003; Loewenstein et al., 2001; Bagozzi et al., 1999). Das aus den Verhaltensabsichten resultierende, konkrete Kundenverhalten hat eine negative Auswirkung auf das Unternehmensergebnis (Rudolf-Sipötz und Tomczak, 2001; Cronin et al., 2000).

Basierend auf dem theoretisch-konzeptionellen Bezugsrahmen wird das provisorische Wirkungsmodell entwickelt. Es handelt sich um die Veranschaulichung eines möglichen (Partial-) Wirkungsmodells, das nach der qualitativen Untersuchung im Rahmen der Dissertation zum zentralen Forschungsmodell erweitert und angepasst wird (vgl. Kap. 4). Die in Abbildung 2-5 dargestellten Ursache-Wirkung-Zusammenhänge zwischen den Modellvariablen werden als Hypothesen formuliert (Bollen, 1989). Sie werden im Rahmen der qualitativen und quantitativen Empirie überprüft. Die grafische Darstellung erfolgt mit Hilfe des sogenannten Pfadmodells (Loehlin, 1987). Die in Kapitel 4.2 aufgestellten Hypothesen sind im Modell als Pfeile eingezeichnet.

Abbildung 2-5: Provisorisches Wirkungsmodell

Legende: B. = Beurteilung, AV = abhängige Variable, UV = unabhängige Variable, Neg. WOM = Negative Word-of-Mouth;

Quelle: Eigene Darstellung

Das provisorische Wirkungsmodell resultiert aus dem S-O-R-basierenden Modell der Wahrnehmungs- und Entscheidungsprozesse von Nolting und Paulus (2005). Es wird zwischen unabhängigen Variablen (vgl. Kap. 4.1.1) unterschieden, die ein Unternehmen beeinflussen kann (Stimulus), und abhängigen Variablen, wie die *aktuellen Prozesse* (Organismus) sowie die *Verhaltensabsichten* (Reaktion).

Im Zentrum des Wirkungsmodells steht die von der Unternehmung initiierte Kundensteuerungsmassnahme. Die Kommunikation (interaktionale Gerechtigkeit) der Push- oder Pull-Massnahme beeinflusst die Kundenwahrnehmung (Smith et al., 1999; Tax et al., 1998; Bies und Shapiro, 1987). Das Modell impliziert, dass eine als Anreiz-Massnahme gestaltete Kundensteuerung kognitiv als fair beurteilt wird (Cropanzano und Folger, 1991; O'Reilly und Puffer, 1989). Setzt das Dienstleistungsunternehmen jedoch Richtlinien und Strafen als Kundensteuerungsmassnahmen ein, resultiert eine tiefe Fairness-Bewertung (Wirtz und Mattila, 2004; Kim, 2003). Die *aktuellen Prozesse*, abgebildet als *Fairness-Bewertung* und *emotionale Reaktion*, sind von der individuellen Disposition (Demografie) geprägt (vgl. Kap. 2.3.1). In der vorliegenden Untersuchung wird der Einfluss der Disposition basierend auf endogenen, (flug-) demografischen Variablen gemessen.

Die Kundenzufriedenheit wird im Modell nicht als eigenständige Variable abgebildet, sondern über die emotionale Reaktion gemessen (Andreassen, 2000; Stauss und Neuhaus, 1997). Da Bestrafungen als Forschungsobjekt im Vordergrund stehen, werden die *negativen* Verhaltensabsichten als Kundenreaktion abgebildet (vgl. Kap. 2.1.4; Kap. 2.4.1).

3 Bestrafung im Airline-Bereich

Für die Erforschung neuer Konstrukte wird häufig eine qualitative Exploration durchgeführt (Malhotra, 2004; Schnell et al., 1999). Um die Forschungsobjekte *Strafe* und *Anreiz-Optionen* näher erfassen zu können, werden sie in diesem Kapitel zuerst anhand der Airline-Branche aus theoretisch-konzeptioneller Sicht bearbeitet (Kap. 3.1). Dazu gehört eine Einführung in die Besonderheiten der Airline-Branche (Kap. 3.1.1 bis Kap. 3.1.3), wobei Massnahmen zur Kundensteuerung bei einer Flugreise im Vordergrund stehen (Kap. 3.1.4). Wie bereits in den theoretisch-konzeptionellen Grundlagen aufgezeigt, spielen zur Abgrenzung von Push-Massnahmen verwandte Konstrukte eine wichtige Rolle. Sie werden in diesem Kapitel aus der Airline-Perspektive erläutert: Kundenzufriedenheit und negative Verhaltensabsichten (Kap. 3.1.5), Dienstleistungszwischenfälle auf Flugreisen (Kap. 3.1.6) sowie die Wiedergutmachung nach einem Dienstleistungszwischenfall (Kap. 3.1.7).

Zur Ergänzung des theoretischen Teils wird das Straf-Konstrukt mit qualitativen Methoden erfasst. Der Untersuchungsfokus (Kap. 3.2) verdeutlicht, dass die eingesetzten qualitativen Instrumente zur Exploration des Straf-Konstrukts und als Vorbereitung zur quantitativen Empirie geeignet sind. Mit Expertengesprächen und einer Inhaltsanalyse wird das provisorische Modell aus der *Anbieterperspektive* überprüft und erweitert (Kap. 3.4; Kap. 3.6). Die *Kundenperspektive* wird mit einer E-Mail-Umfrage und einer Beobachtung berücksichtigt (Kap. 3.3; 3.5). Bevor das Kapitel mit den Folgerungen und Beantwortung der Forschungsfragen abgeschlossen wird (Kap. 3.8), werden die empirischen Grenzen der qualitativen Untersuchung aufgezeigt (Kap. 3.7).

3.1 Grundzüge der Airline-Branche

Wegen hoher Kosten der Dienstleistungserstellung und wegen zunehmendem Wettbewerb sowie tiefer Marktpreise haben sich Bestrafungsrichtlinien in der Airline-Branche als Standard etabliert (Kim, 2006). Es verwundert nicht, dass die Airline-Branche neben dem Finanzdienstleistungssektor zu den Industrien mit den meisten Kundenstrafrichtlinien zählt (Fram und McCarthy, 1999; Bejou und Palmer, 1998). Diese hohe Dichte an Bestrafungsrichtlinien, die Anfälligkeit der Prozesse für Fehler, die von Kunden als Bestrafung wahrgenommen werden, und die Vorreiterrolle bei Kundenanbindungsprogrammen (als *Gegenkonzept* zu Push-Massnahmen, vgl. Kap. 2.1.2) prädestinieren die Airline-Branche zur Untersuchung von Bestrafungen.

3.1.1 Airline-Geschäftsmodelle

Die Airline-Branche wird dem Luftverkehr zugeordnet und als die *Gesamtheit aller Vorgänge* definiert, „die der Ortsveränderung von Personen, Fracht und Post auf den Luftwegen dienen [...] und alle damit unmittelbar verbundenen sonstigen Dienstleistungen“ (Rössger und Hünemann, 1965, S. 3). Der Untersuchungsbereich dieser Arbeit ist begrenzt auf den gewerblichen, öffentlichen und

zivilen Personenluftverkehr. Dieser lässt sich nach den Verkehrsarten *Linien-* und *Gelegenheitsverkehr* sowie nach *Entfernungsgebieten* und *Geschäftsmodellen* unterteilen.

Der Linienverkehr wird im International Civil Aviation Organization (ICAO) Abkommen von Chicago definiert als “any scheduled air service performed by aircraft for the public transport of passengers, mail or cargo” (ICAO, 1944, zit. aus Pompl, 2002). Der Linienverkehr ist netzorientiert und unterscheidet sich dadurch vom Gelegenheitsverkehr. Der Gelegenheitsverkehr, auch Charter-Verkehr genannt, ist vor allem auf Punkt-zu-Punkt-Verkehr ausgerichtet und unterliegt nicht dem Kriterium der Öffentlichkeit⁴ sowie der Tarifpflicht⁵, die als weitere charakteristische Merkmale des Linienverkehrs gelten. Die Abgrenzung nach Entfernungsgebieten kann anhand der Streckenlänge vorgenommen werden (Kurz-, Mittel- und Langstrecken) oder nach Verkehrsregion (Interkontinental-, Kontinental-, Interregional- und Regionalluftverkehr).

Mit *Geschäftsmodellen* wird beschrieben, wie ein Unternehmen, ein Unternehmenssystem oder eine Branche am Markt Werte schafft (Belz und Bieger, 2006). Aus betriebswirtschaftlicher Sicht lassen sich heute im Airline-Bereich vier Geschäftsmodelle identifizieren (Bieger, 2002):

- *Netzwerkgesellschaften*: Oftmals frühere *National Airlines* resp. *Flag Carrier*, die sich seit der Deregulierung weltweit in Allianzsysteme organisieren und durch *Hubbing* (vgl. unten) und optimiertes Netzwerkmanagement einen möglichst grossen Marktanteil zu sichern versuchen;
- *Low Cost Carrier (LCC)*: Anbieter von günstigen Punkt-zu-Punkt-Verbindungen unter Ausnutzung aller möglichen Kostenvorteile im Betrieb, in der Distribution, im Marketing oder in der Finanzierung;
- *Charter Airlines*: Spezialisierung auf Gelegenheitsverkehr, der nicht vom Linienflugverkehr abgedeckt wird, wie z.B. Ferienreisen;
- *Regionalfluggesellschaften*: Fokus auf regionale Nischen, Kostenvorteile und technische Spezialisierung, wie z.B. Bedienung von Flughäfen mit kurzen Pisten.

Diese Geschäftsmodelle haben sich in letzter Zeit aufgeweicht oder es wurden jeweils erfolgreiche Elemente der anderen Geschäftsmodelle eingebaut (Binggeli und Pompeo, 2005). Damit sich Kunden bezüglich der Qualitätsmerkmale orientieren können, wird bei der Buchungsentscheidung zunehmend die Potentialqualität der Airline beurteilt. Einflussfaktoren für die Bewertung sind die Produktionsfaktoren, der Flugplan, das Erscheinungsbild des Personals und das *Image* der Airline (Pompl, 2002). Gerade im Airline-Bereich ist die Marken-Wahrnehmung entscheidend, da sich die Airlines als Dienstleistungsunternehmen schwieriger über Produktattribute differenzieren können

⁴ An einem Linienflug kann ein unbegrenzter Personenkreis teilnehmen, bei Charterflügen hingegen gehört nur eine bestimmte Gruppe dazu, die sich festgelegten Nebenbedingungen unterwirft (Pompl, 2002).

⁵ Die von Airlines veröffentlichten Flugpreise und Beförderungsbedingungen bedürfen für ihre Gültigkeit der Zustimmung einer Erlaubnisbehörde, zumindest für Verbindungen in gewisse Länder (Doganis, 1995).

(vgl. Kap. 3.1.2). Die Marke spielt beim Kauf einer Flugreise, die ein Vertrauensgut und von Passagieren nicht greifbar ist, eine entscheidende Rolle. Passagiere assoziieren bspw. bei traditionellen *Flag Carrier* Attribute des Landes mit der Airline (Kreuzpaintner, 2003). Kunden von *Low Cost Carrier* hingegen achten weniger stark auf die Marke, sondern auf die Attribute *Preis* und *Angebot* (Binggeli und Pompeo, 2005). Für die vorliegende empirische Untersuchung wird das *ex ante Image* der Airline als wichtiger Einflussfaktor im Forschungsmodell identifiziert (vgl. Kap. 2.1.1). Es wird als *ex ante* bezeichnet, weil davon ausgegangen wird, dass sich die Airline-Image-Beurteilung von Kunden im Laufe des Dienstleistungsprozesses ändern kann. Ein Dienstleistungszwischenfall und die mangelhafte Wiedergutmachung können bei Kunden zu Unzufriedenheit mit der Dienstleistung führen und damit zu einer veränderten Image-Wahrnehmung (vgl. Kap. 2.1.3).

Hubs werden heute nicht nur von Netzwerkgesellschaften, sondern auch von *LCC* eingesetzt (Dobruszkes, 2006). Mit Hubs können *Economies of Scope* (Breiteneffekte) ausgenutzt werden, weil mit einer zunehmenden Zahl von Flugstrecken eine quadratisch wachsende Zahl von Flugverbindungen angeboten werden (Shaw, 2002). Zusätzlich werden *Economies of Scale* realisiert, weil auf einen Flughafen konzentriert wird (kostengünstigere Infrastruktur). Schliesslich erlauben Hubs die Dominanz einzelner, wirtschaftsstarker Räume, womit *Economies of Density* realisiert werden (Bieger, 2002). In der jüngeren Vergangenheit haben sich aber Mängel am *Hubbing*-Konzept gezeigt, weil diese anfällig sind für Verspätungen und hohe Kosten verursachen (Love et al., 2006).

3.1.2 Airline-Servicekette

Mit der Airline-Servicekette wird aus kundenorientierter Marketingperspektive aufgezeigt, welche Schritte ein Passagier von der ersten Kundenansprache über Werbung bis zu den Leistungen nach dem Flug durchlaufen muss (Holloway, 2002). Aufgrund des Fokus der vorliegenden Arbeit werden die Stufen (*Online-*) *Reservation und Einkauf*, *Check-In / Boarding* und das Produkt *Flugreise* genauer erläutert (vgl. Abb. 3-1).

Abbildung 3-1: Airline-Servicekette



Quelle: Eigene Darstellung in Anlehnung an Schörcher und Buchholz, 1994

Online-Reservation und -Einkauf

Die *Reservation und der Kauf eines Tickets* hängen vom Vertriebsweg ab. Die Distribution des Airline-Produktes erfolgt über den direkten Vertrieb (Verkaufsbüros, Call Center, Online-Vertrieb,

Automaten) oder über den indirekten Vertrieb (Agenturen, Reiseveranstalter, Online-Vertrieb über Portale etc.). Die Vertriebsstruktur ist zur Zeit im Umbruch, da die Airlines die Zahl der eigenen Vertriebsstrukturen reduzieren und den Ausbau über *Call Center* und Online-Portale aufbauen (Taneja, 2002). In der vorliegenden empirischen Untersuchung steht der Online-Vertrieb an Endkunden im Vordergrund (vgl. Kap. 4.3).

Die Einführung von Internet-Buchungssystemen hat den Airlines pro Buchung zehn bis fünfzehn USD an Kosten eingespart (Myers et al., 2004). Diese Einsparung wird jedoch von der Preistransparenz im Internet marginalisiert (Jayawardhena et al., 2003). Im Durchschnitt hat der Internetkanal die Preise für Flugtickets um 50 bis 100 USD gesenkt (Myers et al., 2004). Heute kann keine Airline auf den Vertriebskanal der Online-Buchung verzichten: *“More than ever, IT is not a luxury but an online necessity. IT is how airlines and airports do business nowadays”* (Coby, 2007, S. 106). Das British Airways (BA) Online-Buchungssystem wurde bspw. im Jahr 2004 bereits von einem Viertel der BA Passagiere genutzt. Im Vergleich zum Jahr 2003 nahm die Anzahl der Online-Buchungen bei BA um 42% zu (o.V., 2004). Aus Sicht der Reisebranche ist die steigende Bedeutung von Online-Buchungen in den USA und in Europa anhand des Wachstums der Marktanteile sichtbar: Im Jahr 2000 wurden in den USA lediglich 4% der Ferienreisen online getätigt, in Europa 1% (o.V., 2001). Im Jahr 2005 wurden in den USA bereits 42% und in Europa 15% der Reisen online gebucht (o.V., 2007a). Gemäss dem Airline IT Trends Survey von Airline Business gilt weltweit für das Jahr 2007: *“Web sales now account for 35% of all ticket sales, and barriers to wider internet use appear to be diminishing [...]”* (Baker, 2007, S. 38). Eine Umfrage der International Air Transport Association (IATA, 2007) bei 10'000 Flugreisenden hat aufgezeigt, dass Flugreisende *“favored more self-service options online and at the departure gate”* (Collis, 2007).

Mit dem Online-Vertrieb wird die Kommunikationspolitik gegenüber dem stationären Agenturvertrieb verbessert und kann von der Airline selbst gestaltet werden: ständige Verfügbarkeit, Multimedialität, Aktualität, Interaktivität, individuelle Kundenansprache und Marktdaten zur Kundennutzung sind die Eigenheiten des Online-Vertriebs (Grob und Ahlert, 2006). Weitere Charakteristika des Online-Vertriebs im Airline-Bereich sind die vollständige Abwicklung von Geschäftsprozessen, hohe Preistransparenz (Preiswettbewerb), Kundensteuerung über Kurzfristvermarktung einzelner Kontingente oder durch Meilengutschriften, Anpassung der Produktpolitik (elektronisches Check-In, Sitzplatzreservierung, Bestellung von Menus, Wunschfilmen etc.) und die Kostenersparnis durch Wegfall der direkten Vertriebskosten (Taneja, 2002; Schögel, 1998).

Die EU-Kommission hat im September 2007 den Online-Buchungsprozess 400 verschiedener Airlines in fünfzehn EU-Ländern untersucht. Die Ermittlungen ergaben, dass über die Hälfte aller Websites Unregelmässigkeiten aufweisen. Die Mängel betreffen die Preisangaben, die Vertragsbedingungen und die Verständlichkeit der übrigen Konditionen (o.V., 2007b). Die betreffenden Unternehmen wurden von den EU-Behörden kontaktiert und aufgefordert, den Inhalt ihrer Internetauftritte gemäss geltendem EU-Recht zu korrigieren oder ihren Standpunkt darzulegen (o.V., 2007b).

Check-In

In dieser Arbeit wird unter *Check-In* der Schalter einer Abfertigungsfirma (*Ground Handling*) als erster persönlicher Kontaktpunkt in der Airline-Servicekette verstanden. Die Abfertigungsfirma fertigt die Passagiere im Namen der Airline ab: Gepäckaufgabe, Sitzplatzreservation, erste Passkontrolle, Sonderwünsche etc. Für den Passagier wird am Check-In deutlich, welche Restriktionen (Kundensteuerung) auf seinem Flugschein gelten, wenn er bspw. einen Upgrade in die nächsthöhere Klasse oder einen bestimmten Sitzplatz reservieren möchte (Holloway, 2002).

Airlines haben mehrere Möglichkeiten beim Management der Stufe Check-In in der Airline-Servicekette. Gemäss Pompl (2002) werden unterschieden: Vorabend-Check-In, Online- oder Automaten-Check-In, Check-In im Hotel oder am Bahnhof, dezentraler Check-In an jedem Schalter der Fluggesellschaft, nach Beförderungsklassen getrennte Schalter und Check-In durch eine Drittpartei (externes Abfertigungsunternehmen oder Airline). Baker (2007) bestätigt auch beim Check-In die zunehmende Bedeutung des Internets. Von hundert untersuchten Airlines setzen im Jahr 2007 bereits 52 Airlines Web (Internet) Check-In ein.

Flugreise

Eine Flugreise wird als *High-Involvement*-Produkt definiert, weil bei allen Kundensegmenten Zeit und Aufwand für die Produktsuche und -auswahl aufgewendet wird (o.V., 2006b). Auch Geschäftsreisende setzen indirekt, bspw. via Assistenz, Zeit und Aufwand für die Produktsuche auf. Die Eigenschaften des Produkts *Flugreise* decken sich mit den charakteristischen Merkmalen von Dienstleistungen (vgl. A.1).

Die Grundleistung des Produkts Flugreise ist die Beförderung einer Person vom Ausgangspunkt zum Zielort (Pompl, 2002). Die Produktdifferenzierung wird durch die Variation des Leistungsniiveaus (Beförderungsklassen) und des Leistungsumfangs bestimmt. Die meisten *Full Service Network Airlines* setzen bspw. die Beförderungsklassen *Economy*, *Business* und *First Class* ein (Holloway, 2002; Shaw, 2002). Die Klassen unterscheiden sich bezüglich des Leistungsumfangs: Flexibilität bei der Reservierung, Gepäcklimiten, Sitzbequemlichkeit, Betreuung sowie Zusatzleistungen an Bord und am Boden, Anrechnung der Meilen im Kundenanbindungsprogramm etc. (Maurer, 2003).

Im Bereich der *On-Board* Dienstleistungen besteht für Airlines angesichts des herrschenden Oligopols in der Passagierflugzeugherstellung ein geringes Differenzierungspotential. Daher sind Serviceleistungen vor, während und nach Beendigung des Fluges sowie der Flugplan (Abflugzeiten, Route, Bedienhäufigkeit) von zentraler Bedeutung (Holloway, 2002).

Viele Kontaktpunkte, die ein Passagier während der Servicekette erlebt, werden nicht von der Airline bedient (vgl. Abb. 3-1) und können zu einer verzerrten Wahrnehmung der Dienstleistungsqualität der Fluggesellschaft führen (Sultan und Simpson, 2000).

3.1.3 Marktsegmentierung nach Flugreisen

Die Marktsegmentierung nach Flugreisen kann gemäss Doganis (2001) aufgeteilt werden in *berufliche Reisen* (Geschäftsreisen, Incentive-Reisen, Kongressreisen, kombinierte Geschäfts- und Urlaubsreisen) sowie *Privatreisen* (Urlaubsreisen, Besuchsreisen und sonstige Privatreisen).

Geschäftsreisende disponieren ihre Flüge oft kurzfristig und erwarten die Möglichkeit der kostenlosen Umbuchung oder Stornierung sowie die Gültigkeit des Flugscheines bei jeder diese Strecke bedienenden Fluggesellschaft (Interlinefähigkeit). Geschäftsreisen unterscheiden sich von Privatreisen durch die höhere Ausgabenbereitschaft und Anforderungen an die Sitzverfügbarkeit, Service und Komfort sowie reibungslosen *Check-In* (Gilbert und Morris, 1995). Innerhalb der Geschäftsreisenden lässt sich eine weitere Segmentierung durchführen (vgl. Tab. 3-1).

Tabelle 3-1: Marktsegmente Geschäftsreisemarkt

Hard Money Traveller	Soft Money Traveller	Medium Money Traveller	Interim Traveller	Periodic Traveller
Unabhängiger Geschäftsmann, der auf eigene Kosten reist.	Firmenangestellter, der auf Kosten der Firma reist.	Konferenz- oder Incentive-Reisender, der innerhalb einer Gruppe reist.	Kombiniert Geschäftsreise mit Privatreise.	Verkäufer, der mehrere stop-overs macht, basierend auf einem regelmässigen Flugplan.

Quelle: IATA (1993)

Der Geschäftsreisende wird aufgrund von Richtlinien (Tarifbedingungen) davon abgehalten, einen preisgünstigen Sondertarif zu buchen, dessen Zielgruppe Privatreisende sind (Pompl, 2002). Richtlinien als Kundensteuerungsmassnahmen sind: Vorausbuchungsfrist, Mindestaufenthaltsdauer oder kostenpflichtige Umbuchung und Stornierung (Holloway, 2002). Obwohl viele grosse Unternehmen zunehmend versuchen, Reisekosten einzusparen, bleiben Geschäftsreisende die wichtigste Kundengruppe, die es zu binden gilt (Pompl, 2002).

Bei *Urlaubsreisen* auf der Langstrecke werden häufig Einschränkungen bezüglich Service und Komfort akzeptiert, wenn dadurch ein tieferer Preis erzielt werden kann (Binggeli und Pompeo, 2005). Sitzverfügbarkeit, Flexibilität oder Flugfrequenz sind von geringerer Bedeutung (Pompl, 2002). Besuchsreisen zu Verwandten oder Freunden werden auch als *Ethnic Travel* bezeichnet. Es ist der Reisestrom zwischen Aus- und Einwanderungsländern oder Herkunfts- und Aufenthaltsländern der Gastarbeiter. Das Nachfrageverhalten ist dem des Geschäftsreisenden ähnlich (Pompl, 2002). Die Marktsegmentierung der Privatreisekunden lässt sich wie folgt zusammenfassen (vgl. Tab. 3-2).

Tabelle 3-2: Marktsegmente Privatreisemarkt

	Preisorientierte Passagiere	Serviceorientierte Passagiere	Flexibilitätsorientierte Passagiere
Beschreibung	<ul style="list-style-type: none"> - Sehen Ticketpreis als wichtigste Determinante; - Flexibilität als notwendige Nebenbedingung. 	<ul style="list-style-type: none"> - Legen Wert auf bequeme Sitze; - Beachten den Preis als Massgrösse; - Mahlzeit- und Zeitschriftenangebot erwartet. 	<ul style="list-style-type: none"> - Verlangen angemessene Umsteige- und Stornierungsmöglichkeiten; - Messen dem Preis gewisse Bedeutung zu.
Anteil	75%	15%	10%

Quelle: Meffert und Bruhn (2000)

3.1.4 Kundensteuerung bei einer Flugreise

Die Nachfrage nach Luftverkehrsleistungen unterliegt je nach Monat, Wochentag oder Tageszeit einer Schwankung. Da die Planung der Angebotsmenge aufgrund der hohen Investitionskosten relativ unelastisch ist, bestehen erhebliche Anforderungen an die Flugplangestaltung. Mit Kundensteuerungsmassnahmen wie Spezialpreisen auf bestimmte Flüge oder Anrechnung einer hohen Anzahl Meilen im Vielfliegerprogramm wird von der Airline versucht, die Schwankungen auszugleichen (Shaw, 2002; Taneja, 2002).

Wie in Kapitel 2.1.2 identifiziert, können Passagiere über Push-Massnahmen (Richtlinien, Strafen) zu leistungskonformem Verhalten motiviert werden. Mit Pull-Massnahmen (Anreize, Optionen) hingegen wird dem Kunden ein höherer Nutzen versprochen, wenn er sich leistungskonform verhält. Im Folgenden werden Push-Massnahmen als *Richtlinien* bezeichnet und deren Auswirkung beim Kunden als *Strafe*. Pull-Massnahmen werden als *Anreize* bezeichnet und deren Ausgestaltung als *Anreiz-Optionen*.

Steuerung der Passagiere über Push-Massnahmen

Bei einer Flugreise entstehen hohe Fixkosten. Jeder leere Sitz oder zu viel Gepäck führen wegen der knapp kalkulierten Ticketpreise zu einem tiefen Deckungsbeitrag pro Passagier (Taneja, 2002). Dies zwingt Airlines zur Einführung von Strafen für Kunden, die sich nicht unternehmens- resp. leistungskonform verhalten. Einige Praktiken wie bspw. das Überbuchen von Flügen können gerechtfertigt werden, weil sich Fluggesellschaften vor Kunden schützen müssen, die ihre Pläne in letzter Minute ändern (Lauer, 2004; Sorell, 1994). In Tabelle 3-3 werden Beispiele von Push- und Pull-Massnahmen im Airline-Bereich aufgeführt. Die in Kapitel 3.6 durchgeführte Inhaltsanalyse der Airline-Internetseiten bildet die Datengrundlage (vgl. A.5.2). Die Massnahmen werden basierend auf dem Airline-Serviceprozess aufgeführt. Die Pull-Massnahmen werden ausschliesslich auf den Airline-Internetseiten angeboten.

Obwohl in der Untersuchung von Fram und Callahan (2001) der Grossteil der Befragten unzufrieden war mit der monetären Bestrafung durch Airlines, hat keiner der Befragten aufgehört, Reisen über dieselbe Airline zu buchen. Es scheint, als habe die Öffentlichkeit die monetären Bestrafungen akzeptiert (Fram und Callahan, 2001). Dies wird einerseits im universellen Charakter der Strafen begründet: fast alle Airlines mit demselben Geschäftsmodell führen dieselben oder ähnliche Bestrafungsmechanismen. Andererseits besteht an Flughäfen meist ein beschränktes Angebot für die gewünschte Strecke und daher wenig Möglichkeit, auf andere Airlines mit weniger oder keinen Bestrafungsrichtlinien auszuweichen. Auf vielen Flugstrecken bestehen Oligopole oder sogar Monopole (Doganis, 2002). Kunden scheinen loyal gegenüber Airlines zu sein, obwohl sie im Grunde unzufrieden sind - neben dem beschränkten Angebot spielen auch artifizielle Einschränkungen wie Bonusprogramme eine Rolle, die Kunden zu „Geiseln“ machen (o.V., 1995). Aber mit der fortschreitenden Deregulierung nimmt auf bestimmten Strecken die Konkurrenz zu, was Kunden die Möglichkeit gibt, auf eine Konkurrenz-Airline umzusteigen, wenn sie unzufrieden sind (Jones und Sasser, 1995).

Tabelle 3-3: Push- und Pull-Massnahmen bei Airlines

	Push-Massnahmen	Pull-Massnahmen
Reservation / Kauf	Gebühr für Umbuchung Kein separater Check-In, Loungezugang, VIP Boarding oder Upgrade Kein Ersatz für verpassten Flug oder verlorenes Ticket Beschränkte Strecken und Plätze für Flüge mit Meilenguthaben	Option zur flexiblen Umbuchung: - VIP Boarding - Lounge-Zugang
Check-In / Boarding	Gebühren für: - Übergepäck - Sperrige Güter - Anzahl Gepäckstücke - Neuausstellung Ticket Limitiertes (Hand-) Gepäck (Grösse, Anzahl)	Option zum Kauf von 15 kg Übergepäck
Flugreise (On-Board)	Keine Sitzplatzreservation Keine Verpflegung	Sitzplatzreservation Verpflegung als Option
Leistungen nach Flug	Limitierte Meilengutschrift	Status im Bonusprogramm

Quelle: Eigene Darstellung, basierend auf qualitativer Empirie (vgl. Kap. 3.8; A.5.2)

Der Markt für Flugreisen wird seit Einzug der *Low Cost Carrier* und der Liberalisierung des Marktes hart umkämpft, was zur Differenzierung über den Preis geführt hat (Doganis, 2002). Eine Flugreise wird immer mehr zu einem Massengut (Nolan et al., 2005). Dies zwingt Flugunternehmen zur strikten Ahndung von Kunden, die Vertragsbedingungen einer Flugreise brechen. Mit dem Trend zu *LCC* haben Kunden begonnen, sich an weniger Dienstleistungen an Bord zu gewöhnen, und die

Erwartungen an einen *LCC*-Flug angepasst (Binggeli und Pompeo, 2005; de Coverly et al., 2002).⁶ Gerade *Low Cost Carrier* nutzen die Gebühren als Möglichkeit, Zusatzumsätze zu generieren: “*Not only have the no-frills airlines ratcheted up their prices as routes get more popular, passengers are stung for extra charges for overweight bags or checking in late*” (o.V., 2002b).

Die empirische Untersuchung von Fram und Callahan (2001) hat gezeigt, dass es Flugpassagieren selten gelingt, bei Fluggesellschaften eine angedrohte Strafzahlung zu verhindern. Der Grund dafür liegt in der Schwierigkeit der Unternehmung, Gründe für den Vertragsbruch des Kunden zu evaluieren und Strafminderung bei unvermeidbaren Situationen zu gewähren. Dennoch würde sich für Airlines die Aufhebung von gewissen Bestrafungen lohnen. In Industrien, wo Bestrafungen zur Industrienorm gehören, können sich Unternehmen durch die Einführung neuer Kundensteuerungsmassnahmen für spezifische Kundensegmente erfolgreich positionieren (Kim und Smith, 2005). Hierfür müssen die Kundensegmente identifiziert werden, welche bspw. ohne Strafandrohung mit einer hohen Wahrscheinlichkeit die Flugreise antreten werden (McCarthy und Fram, 2000).

Steuerung der Passagiere über Pull-Massnahmen

Airlines nehmen eine Vorreiterrolle bei der Anreizsteuerung über Bonusprogramme ein (Lauer, 2004). Ein Vielfliegerprogramm (*Frequent Flyer Program*) ist ein Bonussystem einer Airline, bei dem der Passagier für einen durchgeführten Flug eine bestimmte Anzahl von Bonuspunkten erhält (Holloway, 2002). Der Passagier kann die Bonuspunkte gegen eine Prämie eintauschen. Hauptziel des Vielfliegerprogramms ist die Bindung des Kunden an die Airline und die Gewinnung von Kundeninformationen über die teilnehmenden Passagiere (Shaw, 2002). Mit dem Vielfliegerprogramm können zudem weniger gut ausgelastete Flüge über Anreize beim Meilen sammeln oder einlösen attraktiv gestaltet werden (Pompl, 2002). Das System zur Anrechnung von Bonuspunkten ist jedoch kompliziert und schwer durchschaubar, was Passagiere verärgert (Imhof, 2006).

Airlines sind besonders dafür geeignet, Anreize wie Service-Upgrades oder Zugang zur Lounge als Bonusleistung und ökonomischen Nutzen anzubieten (Klein, 1998). Die „Produktnutzung“ zeichnet sich durch einen hohen Grad an direkten Kontakten vom Kunden mit der Airline aus. Daher können die verschiedenen Kundensteuerungsmassnahmen mit Anreizwirkung direkt beim Kunden umgesetzt werden (vgl. Kap. 3.1.2).

Status und Exklusivität sind weitere wichtige Komponenten von einem Vielfliegerprogramm (vgl. Kap. 2.1.2). Durch die Klassifizierung des Produktangebots (*Economy*, *Business* und *First Class*) haben Passagiere gelernt, in verschiedenen Statuskategorien zu denken (Lauer, 2004). Bei der Inanspruchnahme der Leistung sind zudem andere Kunden zugegen, so dass der Status für diese sichtbar

⁶ Trotzdem sind Passagiere nicht bereit, auf jegliche Zusatzdienstleistung bei *LCC* zu verzichten: Im Juli 2002 haben mehr als 130 Passagiere ihren Platz im Flugzeug von Easy-Jet nicht verlassen, nachdem sie aufgefordert wurden, ihren Weiterflug zu unterbrechen, um einer gestrandeten Easy-Jet Gruppe-Platz zu machen (o.V., 2002a). *LCC* wie Ryanair offerieren zwar gestrandeten Passagieren einen Platz in der nächsten Maschine oder Rückerstattung, aber viele Passagiere müssen mit erheblichen Unannehmlichkeiten rechnen wie bspw. Übernachtungen im Flughafen oder fehlender Unterstützung bei der Suche nach Hotels (o.V., 2005a).

gemacht werden kann. Um das Verhalten von Kunden zu steuern, reicht bereits die Möglichkeit zum Staterwerb aus und nicht ausschliesslich der Statusgewinn (Lauer, 2004). Dass Status oft wichtiger ist als die tangiblen Leistungen im Bonusprogramm zeigt sich am Vielfliegerprogramm von *Continental Airlines*. 1995 erhöhte Continental die Schwelle zum Erhalt eines Freifluges sowie die Punkte für einen Status-Upgrade. Diese Massnahme löste Proteste bei Vielfliegern aus, worauf sich Continental entschuldigte. Die Statusschwelle für einen Upgrade wurde herabgesetzt. Dass die höhere Schwelle für den Freiflug erhalten blieb, störte die Passagiere nicht (Butscher, 1998).

Ein wichtiger Teil bei der Kundensteuerung über Anreize ist die *Kommunikation*, welche bei Airlines vorzugsweise über das Internet oder über den Versand von Kunden-Zeitschriften erfolgt (vgl. Kap. 3.1.2; Pompl, 2002). Wie die qualitativ-empirische Untersuchung zeigt, stellt vor allem das Internet ein effizientes System zur Kundensteuerung dar (vgl. Kap. 3.6).

3.1.5 Kundenzufriedenheit und negative Verhaltensabsichten

Beim Automobilkauf besteht ein enger Zusammenhang zwischen Kundenzufriedenheit und Loyalität (Jones und Sasser, 1995). Im Gegensatz dazu wirkt sich im Airline-Bereich eine Erhöhung der Kundenzufriedenheit nur bei *unzufriedenen* Kunden auf die Loyalität aus (Söderlund, 1998; Zeit-haml et al., 1996). Kunden, die bereits sehr zufrieden sind, werden durch Massnahmen zur Erhöhung der Zufriedenheit nicht loyaler. Dies wird von Jones und Sasser (1995) dahingehend interpretiert, dass es im Airline-Bereich nur relativ wenige Alternativen auf einer Strecke gibt. Wenn viele Alternativen bestehen, wie im Automobilbereich, offeriert Loyalität eine Lösung für die Multioptionalität und die daraus resultierende Entscheidungsschwierigkeit (Gross, 2005). Auf einem Markt mit wenigen Alternativen ist Loyalität für unzufriedene Kunden eine Lösung auf das Problem der Alternativensuche. Daher ist die Verknüpfung Zufriedenheit-Loyalität schwach, wenn Kunden sehr zufrieden sind, und stark, wenn Kunden unzufrieden sind. Auch das Paradox, dass Kunden, die sich beschwerten, am loyalsten sind, wird mit den fehlenden Alternativen erklärt (Hays und Hill, 1999).

Die Zufriedenheit mit einer Dienstleistung liegt nicht auf einem Kontinuum zwischen unzufrieden und zufrieden (Price et al., 1995). Eine Flugreise kann beim Kunden gleichzeitig Zufriedenheit über das gute Essen an Bord verursachen und negative Gefühle über die Enge der Kabine auslösen (Arnould und Price, 1993). In einem Feldversuch (Taylor, 1994) sowie in einer Laboruntersuchung (Dubé et al., 1996) wurde nachgewiesen, dass Passagiere nach einem Flug mit der Dienstleistung zufrieden waren, obwohl das Flugzeug verspätet abflog.

Bei *High-Involvement*-Produkten wie einer Flugreise ist die Neigung zur Beschwerdeeinreichung höher (Singh und Wilkes, 1996). Weiser (1995) hat das Beschwerdeverhalten nach einem Dienstleistungszwischenfall bei einer Flugreise (vgl. nächstes Kapitel) untersucht und herausgefunden, dass 50% der Kunden, die sich entschlossen haben, der Airline nichts über das negative kritische Ereignis mitzuteilen, den Anbieter gewechselt haben (auf den Strecken, wo es Alternativen gibt).

87% derjenigen, die sich beschwerten und Wiedergutmachung erhielten, blieben der Airline gegenüber loyal. Bamford und Xystouri (2005) untersuchten das Beschwerdeverhalten von Passagieren einer internationalen Airline. Die meisten Beschwerden konnten den folgenden drei Kategorien zugeordnet werden: (1) Verspätungen, meist technischer Natur; (2) Unterbrechung der Dienstleistungserstellung, z.B. durch Streiks und (3) Verhalten vom Bodenpersonal.

Bei der untersuchten Airline gab es pro 1'500 Passagieren im Jahr 2000 eine Beschwerde. Von insgesamt 870 Beschwerden wurden 626 monetär abgegolten und ein Total von 138'904 USD an Kompensationszahlungen geleistet (vgl. Tab. 3-4). Deutlich erkennbar ist in dieser Untersuchung, dass die Totalvergütungen überproportional und die Durchschnittskosten für Beschwerden zwischen 1997 und 2000 um 30% gestiegen sind.

Tabelle 3-4: Passagierbeschwerden und monetäre Vergütung

	Anzahl Passagiere	Anzahl Beschwerden	Anzahl finanziell vergütet	Total vergütet in USD	Ø-kosten pro Beschwerde
1997	1'298'000	545	253	\$42'573	\$168
1998	1'355'000	684	403	\$88'090	\$218
1999	1'346'000	715	462	\$131'534	\$285
2000	1'400'000	870	626	\$138'904	\$222
Veränderung 97 - 00	x 1,08	x 1,6	x 2,5	x 3,3	x 1,3

Quelle: Eigene Darstellung in Anlehnung an Bamford und Xystouri, 2005

3.1.6 Dienstleistungszwischenfälle auf Flugreisen

Bei einer Flugreise besteht eine hohe Anfälligkeit für Störungen resp. negative kritische Ereignisse im Serviceprozess (Verma, 2001; Bejou und Palmer, 1998). Folgende Faktoren werden von Flugreisenden als Dienstleistungszwischenfall angegeben: Verspätungen, Überbuchungen, unbequeme Sitze und laute Passagiere in der Kabine (de Coverly et al., 2002; Gilbert und Morris, 1995; Edvardsson, 1992). Verma (2001) hat in einer empirischen Untersuchung bei Flugreisenden vier Arten von Dienstleistungszwischenfällen im Airline-Bereich identifiziert:

- *Baggage Handling* (verlorene, nicht angekommene oder beschädigte Koffer);
- *In-Flight Care and Services* (Crew, Verpflegung, Unterhaltung, Sitzplatz, Ruhe);
- Humanfaktoren (Auskünfte, Hilfe bei Problemen, Verspätungen etc.);
- Sonstige Probleme (Technik, Reservation, Ansagen etc.).

Die meistgenannten Zwischenfälle, die beobachtbar waren, gehörten zur Gruppe von *In-Flight Care and Services*. Insbesondere waren die von Verma befragten Kunden unzufrieden mit der Qualität des Essens und dem Angebot an Lesematerial. Gegenüber den anderen Zwischenfällen waren Flugreisende erstaunlicherweise indifferent. Unzufrieden waren Kunden mit den Bemühungen von Airlines zur Wiedergutmachung. Die empirische Untersuchung mit 1'956 Passagieren von Sultan und Simpson (2000) bestätigt die Erkenntnis, dass Passagiere Fehler im Airline-Serviceprozess akzeptieren, wenn eine Wiedergutmachung erfolgt. Zudem wird aufgezeigt, dass für Kunden die Zuverlässigkeit der Airline der wichtigste Aspekt in der Beurteilung der Dienstleistungsqualität ist.

Bitner et al. (1990) bestätigen den Stellenwert der Wiedergutmachung. Sie haben 700 negative kritische Ereignisse im Dienstleistungsbereich untersucht und herausgefunden, dass es nicht das Ereignis selbst ist, das zu Unzufriedenheit führt - die meisten Kunden akzeptieren, dass etwas schief gehen kann. Es ist vielmehr die Reaktion der Unternehmen auf ein kritisches Ereignis, welches zu Ärger und Unzufriedenheit führt: "*Good recovery programs can turn angry customers into loyal ones and are vital in complex services, like airlines, that have many factors outside of their control*" (de Coverly et al., 2002, S. 25).

3.1.7 Wiedergutmachung nach Dienstleistungszwischenfall

Wiedergutmachung (*Service Recovery*) beinhaltet die Schritte, die ein Dienstleistungsunternehmen durchführt, um auf einen Dienstleistungszwischenfall zu reagieren (Johnston und Mehra, 2002; Grönroos, 2000). Es fehlt aber bei vielen Airlines an effizienten Massnahmen zur Wiedergutmachung: "*It may be due to lack of appreciation of the importance of achieving recovery in service businesses by the top management*" (Verma, 2001, S. 77). Gerade im Airline-Bereich, wo auf einzelnen Strecken kein Wettbewerb herrscht, wird dies augenfällig.

Strategien der Wiedergutmachung umfassen drei verschiedene Aktionen: *Entschuldigung*, *Unterstützung* und *Kompensation* (Levesque und McDougall, 2000; Boshoff, 1997; Hart et al., 1990). Unterstützung heisst bspw., dass die Airline bei einem abgesagten Flug dem Kunden hilft, einen anderen Flug oder Hotelübernachtungen zu buchen. Eine Entschuldigung sowie Unterstützung reichen nicht unbedingt; Kunden erwarten teilweise ein *Upgrade* oder Gutscheine für weitere Reisen (de Coverly et al., 2002). Zudem ist das *Timing* sehr wichtig: je schneller reagiert wird, umso höher ist die Kundenzufriedenheit (Boshoff, 1999).

United Airlines (UA) hat eine gute Balance gefunden zwischen Entschuldigung, Kompensation und *Timing*. Weil viele Flüge wegen schlechten Wetters und streikender Mitarbeiter abgesagt werden mussten, verlor UA wichtige Kunden. UA reagierte darauf mit einer Werbesendung, wo sich CEO James Goodwin für die Probleme der Airline entschuldigte, diese erläuterte und Kompensation in Form von Gebührenreduktionen anbot (Strout, 2000). Dies war eine vorbildliche Wiedergutmachung, weil sich die Airline entschuldigte, Wiedergutmachung anbot und die Kunden innert nützlicher Frist über die Gründe für das Problem aufgeklärt hat (de Coverly et al., 2002).

Boshoff (1997) hat 540 Passagiere befragt und dabei mit einem Dienstleistungszwischenfall konfrontiert (ein verpasster Anschlussflug wegen verspäteter Ankunft). Bei einer negativen Dienstleistungserfahrung sind gemäss dieser Untersuchung folgende drei Strategien zur Wiedergutmachung erfolgversprechend: (1) Eine schnelle Antwort eines hochrangigen Angestellten; (2) eine schnelle Antwort eines Angestellten mit vollständiger Kompensation oder (3) ein hoher Unkostenbeitrag, der von einem hochrangigen Angestellten übergeben wird. Eine Entschuldigung allein genügt nicht als erfolgreiche Handlung, wenn sie nicht begleitet ist von einer tangiblen Kompensation (Boshoff, 1997).

In Situationen hingegen, wo sich die Unternehmung schnell entschuldigt *und* Rückerstattung leistet, erhöhen zusätzliche Kompensationszahlungen die Kundenzufriedenheit nicht (Wirtz und Mattila, 2004). Daher ist es eine effiziente Strategie, sich so schnell als möglich zu entschuldigen sowie den Schaden zu beheben, ohne dabei zusätzlich zum Kostenbeitrag eine Kompensationszahlung leisten zu müssen. Es braucht im Sinne der *Justice Theory* interaktive und prozedurale Gerechtigkeit (vgl. Kap. 2.3.2), um die Kundenzufriedenheit zu gewähren, wenn keine Kompensation geleistet wird. Dies gilt vor allem, wenn Kunden glauben, die Unternehmung habe Kontrolle über die Situation gehabt (Taylor, 1994). Verspätete Antworten auf einen kritischen Zwischenfall lösen bei Kunden mit hoher Wahrscheinlichkeit Wut als emotionale Reaktion aus. Eine verspätete Antwort impliziert eine ineffiziente Organisation und daher eine hohe Wahrscheinlichkeit aus Sicht des Kunden, dass sich in Zukunft wieder ein Zwischenfall ereignet (Wirtz und Mattila, 2004).

3.2 Untersuchungsfokus der qualitativen Empirie

Der Untersuchungskontext der qualitativen Empirie wird in Kapitel 3.2.1 erläutert. Ebenfalls zum Untersuchungsfokus gehören das Erkenntnis- und Forschungsobjekt (Kap. 3.2.2) sowie das Forschungsziel (Kap. 3.2.3) der qualitativen Empirie. Die Begründung der Methodenauswahl für jeden einzelnen Schritt der qualitativen Empirie wird in Kapitel 3.2.4 vorgenommen.

3.2.1 Untersuchungskontext

Der Untersuchungskontext der vorliegenden qualitativen Empirie sind der *Flugpassagier und die Steuerung seines Verhaltens in der Airline-Servicekette* über Strafen oder Anreiz-Optionen. Verschiedene kritische Ereignisse im Airline-Serviceprozess wurden bereits erforscht (vgl. Kap. 3.1.6). Inwiefern diese Ereignisse von Flugpassagieren als Strafe wahrgenommen werden und wie die Push-Massnahmen mit Pull-Massnahmen als „Gegenkonzept“ kombiniert werden können, ist unklar. Insbesondere emotionale und kognitive Kundenreaktionen auf (Straf-) Ereignisse wurden bisher nicht erforscht. Diese Lücke wird mit der vorliegenden, qualitativen Untersuchung geschlossen (vgl. Kap. 1.3).

3.2.2 Erkenntnis- und Forschungsobjekt

Das Erkenntnisobjekt sind das provisorische Wirkungsmodell zur Kundenwahrnehmung und zum Kundenverhalten sowie die dazugehörigen Hypothesen (vgl. Kap. 2.6; 4.2). Im Vordergrund stehen negative Ereignisse, die bei Flugpassagieren als Strafe wahrgenommen werden, sowie Anreizsysteme als alternative Kundensteuerung. Die qualitative Empirie stützt sich somit auf Erkenntnisse aus dem Dienstleistungsbereich zur Kundensteuerung und zu negativen Verhaltensabsichten (vgl. Kap. 2.1), der (psychologischen) Wahrnehmungs- und Entscheidungsprozesse von Individuen (vgl. Kap. 2.3) und der Forschung zum Airline-Bereich aus Kapitel 3.1 (Servicekette, Geschäftsmodelle etc.).

3.2.3 Forschungsziel

Das Ziel der qualitativen Untersuchung ist die Beantwortung der Forschungsfragen b.1 und b.2 (vgl. Kap. 1.4). Einflussfaktoren auf die Wahrnehmung und Verhaltensabsichten verschiedener Kundensegmente sollen erarbeitet werden (b.1). Zudem wird eine Konzeptualisierung der Kundensteuerungsmassnahmen *Strafe* und *Anreiz-Option* im Airline-Bereich durchgeführt. Es soll beantwortet werden, wie und auf welcher Stufe Airlines ihre Passagiere steuern (b.2). Neben der Beantwortung der Forschungsfragen wird neuer Input für das zu erarbeitende Forschungsmodell generiert. Einerseits werden für die quantitative Untersuchung die inhaltlichen Hypothesen erweitert und angepasst. Andererseits sollen die Grundlagen zur Fragebogenentwicklung erarbeitet werden (Variablenselektion, Operationalisierung der Variablen). Die Forschungsfragen werden in Kapitel 3.8 beantwortet.

3.2.4 Methodenauswahl

Yin (1994) und Göthlich (2003) empfehlen bei qualitativer Forschung die Durchführung einer Triangulation: Aussagen werden von mehreren unabhängigen Perspektiven betrachtet und auf ihre Statthaftigkeit überprüft. Für die vorliegende, qualitative Untersuchung werden daher verschiedene Erhebungsmethoden eingesetzt. Das Konstrukt *Bestrafung im Airline-Bereich* sollte zudem aus Kunden- und Anbietersicht untersucht werden. Um diese Kunden- und Anbietersicht abzubilden, die Anforderungen an die Triangulation zu erfüllen und die Stärken sowie Schwächen verschiedener qualitativer Erhebungsmethoden auszugleichen, erfolgt die Datenerhebung in vier sequentiellen Schritten (vgl. Abb. 3-2):

1. Exploratorische E-Mail-Umfrage bei Vielfliegern (Kundenperspektive);
2. Leitfadenterviews mit Experten aus verschiedenen Bereichen im Airline-Serviceprozess (Anbieterperspektive);
3. Beobachtung am Check-In (Kundenperspektive);
4. Inhaltsanalyse des Internet-Buchungsprozesses verschiedener Airline Websites (Anbieterperspektive).

Die Reihenfolge der Erhebungen ergibt sich aus den Anforderungen an das analytische Raster. Ergebnisse der exploratorischen E-Mail Umfrage dienen der Erarbeitung des Leitfadenterviews, die Expertengespräche helfen beim Aufbau des Kategoriensystems für die Beobachtung (Identifikation kritischer Ereignisse am Check-In) und Erkenntnisse aus der Beobachtung dienen zur Entwicklung der Online-Inhaltsanalyse.

Abbildung 3-2: Aufbau und Durchführung der qualitativen Empirie

	Forschungs- und Erkenntnisobjekt	Forschungsziele	Erhebungs- und Auswertungsmethodik
E-Mail-Umfrage	<ul style="list-style-type: none"> – Vielflieger (Kundensicht) – Strafen und Kundenverhalten 	<ul style="list-style-type: none"> – Evaluation von Bestrafungs-Szenarien und negative Verhaltensabsichten – Fairness als Indikator 	<ul style="list-style-type: none"> – E-Mail-Versand an Vielflieger ($N = 31$) – Deskriptive Datenanalyse
Experten-interview	<ul style="list-style-type: none"> – Kundenreaktionen (Anbietersicht) – Kundensteuerung im Airline-Bereich 	<ul style="list-style-type: none"> – Überprüfung und Adaption Hypothesen, Szenarien – Anreiz-Steuerung Internet – Selektion, Op. Variablen 	<ul style="list-style-type: none"> – Leitfadengespräch mit Airline-Experten und Abfertigungsfirmen ($N = 11$) – Inhalts- und Dimensionsanalyse
Beobachtung	<ul style="list-style-type: none"> – Passagiere in Zürich auf verschiedenen Airlines und Strecken – Provisorisches Wirkungsmodell 	<ul style="list-style-type: none"> – Check-In als Service-Stufe, Kommunikation und Ticketpreis als UV – Kundenwahrnehmung (Emotion) und Verhalten 	<ul style="list-style-type: none"> – Beobachtung am Check-In Zürich bei Swissport und Jet Aviation ($N = 478$) – Inhaltsanalyse
Inhalts-analyse	<ul style="list-style-type: none"> – Airline-Internet-Buchungsprozesse – Vorbereitung quant. E. 	<ul style="list-style-type: none"> – Preise Ticket, Optionen – Kommunikation, optische Darstellung auf Website 	<ul style="list-style-type: none"> – Zugriff auf Websites ($N = 12$) – Inhaltsanalyse

Legende: Op. = Operationalisierung, N = Gesamtstichprobe, quant. E. = quantitative Empirie;
Quelle: Eigene Darstellung

In den folgenden Kapiteln 3.3 bis 3.6 werden das Erkenntnis- und Forschungsobjekt, das Forschungsziel, die Datenauswertung und -erhebung sowie die Forschungsergebnisse jedes einzelnen Schrittes der qualitativen Empirie aufgeführt.

3.3 E-Mail-Umfrage bei Vielfliegern

Als erster Schritt der qualitativen Empirie wurde im Juni 2006 eine *exploratorische E-Mail-Umfrage* unter Vielfliegern durchgeführt. Für die theoretischen Grundlagen zur Forschungsmethode *Befragung* wird auf den quantitativ-theoretischen Teil verwiesen (vgl. Kap. 5.2).

3.3.1 Erkenntnis- und Forschungsobjekt

Die exploratorische E-Mail-Umfrage beim Forschungsobjekt *Vielflieger* wird als erster Schritt durchgeführt, um möglichst ökonomisch zu ersten Erkenntnissen bezüglich Strafsituationen und

negativen Verhaltensabsichten im Airline-Bereich zu gelangen. Es wird angenommen, dass Vielflieger über Erfahrungen oder Wissen bezüglich Strafsituationen verfügen, da Strafmassnahmen in der Airline-Branche verbreitet sind (McCarthy und Fram, 2000; Bejou und Palmer, 1998).

3.3.2 Forschungsziel

Mit der exploratorischen E-Mail-Umfrage werden Strafsituationen bezüglich der Verhaltensabsichten evaluiert und neue Szenarien für den Fragebogen generiert. Zudem wird untersucht, ob die Fairness-Beurteilung ein geeigneter Indikator für die Bewertung der Kundensteuerungsmassnahmen ist.

3.3.3 Datenerhebung und -auswertung

Es wurden 102 Personen aus privaten E-Mail-Datenbanken kontaktiert. 31 Personen haben den Fragebogen beantwortet. 24% waren weiblich; das Durchschnittsalter betrug 34 Jahre und 86% der Befragten fliegen mehr als viermal pro Jahr (Zürich als Abflugort), womit die Stichproben-Voraussetzung „Vielflieger“ erfüllt wurde.

Die exploratorische E-Mail-Umfrage wird mit einer deskriptiven Datenanalyse ausgewertet. Die Ergebnisse der Fairness-Beurteilung sind auf einer 5er Likert-Skala (vgl. Kap. 5.2) als Durchschnittswerte abgebildet. Die Skala zur Fairness-Beurteilung wurde von Schoefer und Ennew (2005) adaptiert. Offene Fragen ermöglichen zusätzlichen Input zu Strafsituationen (vgl. A.4.1).

3.3.4 Forschungsergebnisse

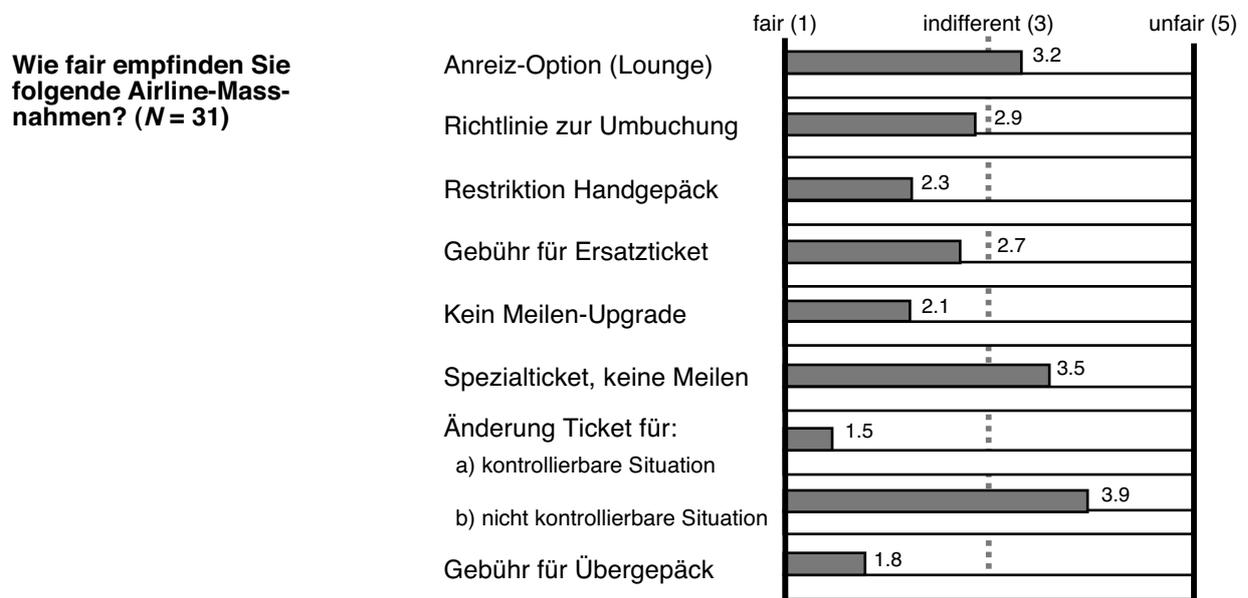
Die Ergebnisse der E-Mail-Umfrage bestätigen, dass Kunden vor allem Situationen, die sie nicht unter Kontrolle hatten, als unfair empfinden (vgl. Abb. 3-3). Insbesondere ein nicht gewährter Meilen-Upgrade führt zu einer negativen Fairness-Beurteilung. Der Meilen-Upgrade wird als valides Szenario für den Fragebogen der quantitativen Untersuchung bestätigt. Die Einstufung der Übergepäckgebühr als *faire* Strafe zeigt aus Kundenperspektive die verbreitete Akzeptanz und das Verständnis dafür (vgl. Kap. 1.2). Trotzdem wird das Übergepäck-Szenario in der quantitativen Untersuchung verwendet, da die empirische Beobachtung emotionale Reaktionen bei Übergepäck und negatives Verhalten bestätigt hat (vgl. Kap. 3.5).

Die verwendete Skala zur Fairness-Wahrnehmung hat sich bewährt. Sie wird für die quantitative Untersuchung verwendet. Die Fairness-Beurteilung als Indikator für die Kundensteuerung (Richtlinie oder Anreiz) wird bestätigt.

Die meist genannten Bestrafungen, die die Befragten bereits erlebt haben (und als *Bestrafung* deklarierten), sind: Verspätungen, Überbuchung der Sitzplätze, Kosten für Übergepäck und Umbuchung sowie kompliziertes Check-In. Fast die Hälfte der Befragten (42%) waren überrascht, dass sie bestraft wurden. Sie nannten als Grund die mangelhafte Kommunikation durch die Airline. Negative

Mund-zu-Mund-Propaganda wurde als häufigste Reaktion auf den Zwischenfall genannt (67%). Obwohl sich drei Viertel der Befragten nach dem Zwischenfall frustriert fühlten, beschwerte sich nur eine Person bei der Airline (vgl. A.4.1 für den Fragebogen).

Abbildung 3-3: Kundenbefragung zur Fairness-Beurteilung von Richtlinien



Legende: FFP = *Frequent Flyer Program*

Quelle: Eigene Darstellung

3.4 Leitfadeninterview mit Experten

Die Expertengespräche werden mit einem teil-strukturierten (*Leitfaden-*) *Interview* durchgeführt. Für die Erfassung kognitiver Vorgänge, die aufgrund ihrer Einmaligkeit, Prozesshaftigkeit und Situationsspezifität mit anderen Instrumenten kaum zugänglich sind (vgl. Kap. 2.3.1), erscheint die Befragungsmethode mit einem Leitfadeninterview vorteilhaft (Lamnek, 2005). Da die Kundensteuerung über kognitive Wahrnehmungsprozesse bestimmt wird (vgl. Kap. 2.6), eignet sich das Leitfadeninterview zur Beantwortung der Forschungsfragen (b.1, b.2). Dabei wird ein Fragebogen verwendet, bei dem der Interviewer Wahlfreiheit bezüglich der Frageabfolge und Formulierung hat (Stier, 1999). Der Interviewer stellt beim Leitfadengespräch die Fragen gemäss dem sich entwickelnden Gesprächsverlauf (Latz, 1993).

3.4.1 Erkenntnis- und Forschungsobjekt

Aufbauend auf dem theoretisch-konzeptionellen Teil wird das Leitfadeninterview mit Experten von Airlines und vom Flughafen durchgeführt. Das Forschungsobjekt ist die *Kundenreaktion auf negative kritische Ereignisse* und *Kundensteuerung im Airline-Bereich*. Die Thematik wird aus der Anbieterperspektive untersucht.

3.4.2 Forschungsziel

Ziel der Expertengespräche ist, die Thematik tiefer zu durchdringen und bestehende Hypothesen auf ihre Praxisnähe zu untersuchen und zu erweitern. Daneben dienen die Gespräche zur Selektion und Operationalisierung der Variablen des zu erarbeitenden Forschungsmodells. Eine Auswahl an Szenarien für die quantitative Untersuchung wurde aus Erkenntnissen des theoretischen Teils (vgl. Kap. 3.1) erstellt. Mit den Expertengesprächen soll überprüft werden, ob sich die Szenarien als Massnahmen zur Kundensteuerung eignen. Ziel ist, vier praxisnahe Szenarien, die sich als Anreiz-Option und als Bestrafungsrichtlinie gestalten lassen, zu definieren. Mit einer Dimensionsanalyse zum Konstrukt *Bestrafung im Airline-Bereich* werden die Ergebnisse der Expertengespräche mit Erkenntnissen des theoretischen Teils kombiniert (vgl. Abb. 3-4).

3.4.3 Analytisches Raster zur Datenauswertung

Im Sinne von Leitfadeninterviews werden die Expertengespräche mit vorbereiteten Themengebieten in einem offenen Gespräch geführt. Zur Befragung gehört, dass „Bestrafung von Kunden“ neutral ausgesprochen wird (z.B. „unerwünschte Konsequenzen für den Kunden“). Jedes Expertengespräch ist wie folgt strukturiert:

- Erläuterung des provisorischen Wirkungsmodells und der Hypothesen;
- Beurteilung potentieller (zusätzlicher) Einflussfaktoren auf die Fairness-Beurteilung;
- Entwicklung von Szenarien für die Fragebogenentwicklung. Praxisrelevanz und Diskussion der im Szenario verwendeten Preise;
- Diskussion über geeignete Stufen im Airline-Serviceprozess zur Untersuchung der Kundensteuerung über Anreize oder Richtlinien;
- Fragen zu den negativen Verhaltensabsichten (Beschwerden) der Passagiere.

3.4.4 Datenerhebung

Befragt wurden elf verschiedene Experten in den Bereichen *Customer Care*, *Passenger Services* sowie *Market Research*, *Marketing* und *E-Commerce*. Die ausgewählten Airline-Experten sind Mitarbeiter von Swiss und Singapore Airlines. Bei diesen Airlines handelt es sich um Firmen mit hohen Qualitätsansprüchen gegenüber Kunden (vgl. Sobie, 2007). Die Anbieterperspektive im Airline-Serviceprozess umfasst neben den Airlines auch den Flughafen. Daher wurden Experten der *Ground-Handling*-Firmen Jet Aviation und Swissport sowie vom Flughafen Zürich befragt. Im Anhang sind die Experten und die Interviewdaten aufgeführt (A.6).

3.4.5 Forschungsergebnisse

Steuerung von Passagierverhalten

Die Expertengespräche haben verdeutlicht, dass die Kundensteuerung mit online angebotenen Optionen von Interesse und umsetzbar sind. Die Kommunikation der Anreiz-Massnahmen spielt dabei eine zentrale Rolle, wie Erfahrungen aus dem Vielfliegerprogramm zeigen. Mit dem Vielfliegerprogramm werden Passagiere heute aktiv und erfolgreich von Airlines gesteuert. Als quantitativ zu analysierende Anreiz-Optionen werden ein käuflicher Lounge-Zugang, VIP Boarding am Gate (Boarding vor den anderen Passagieren) sowie eine deutlichere, web-unterstützte Kommunikation eines flexiblen Tickets bestimmt. Der Preis der Anreiz-Optionen wird mit der Inhaltsanalyse für den Fragebogen der quantitativen Untersuchung ermittelt (vgl. A.5.2).

Die empirischen Ergebnisse der Kundenbefragung zu kritischen Ereignissen von Verma (2001) werden mit den Erkenntnissen der Experteninterviews aus der Airline-Perspektive validiert. Basierend auf dem Beschwerdeverhalten von Kunden wurden folgende kritischen Ereignisse identifiziert:

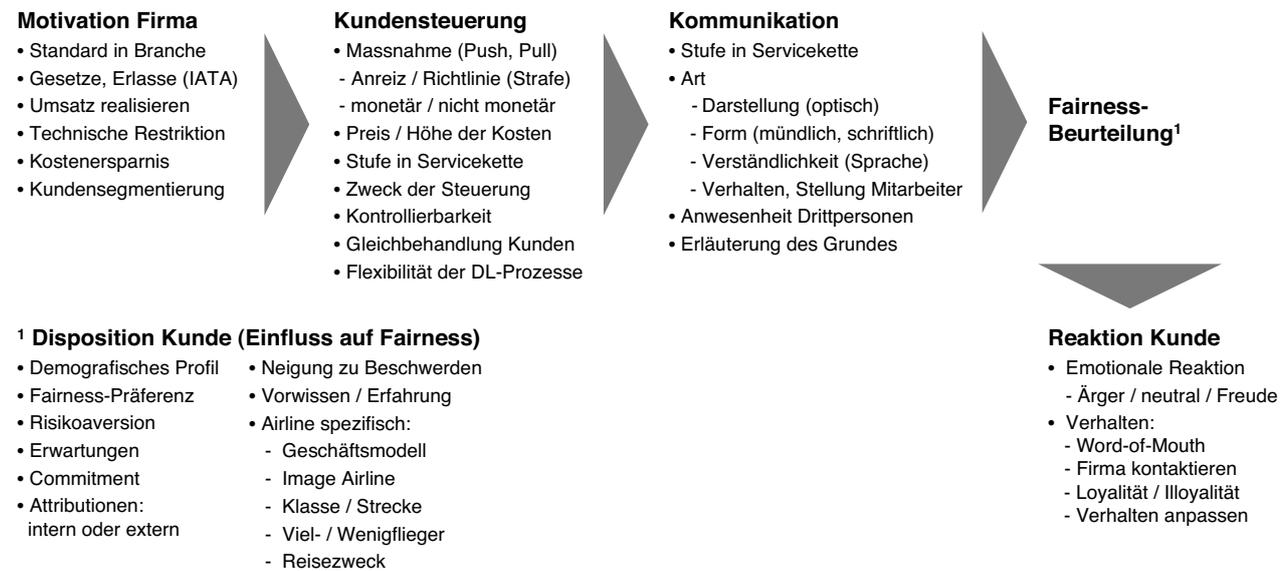
- Mangelndes Wissen von Kunden bezüglich der Tarifbedingungen, insbesondere: Übergepäckgebühr, Anzahl Gepäckstücke am Check-In, Grösse des Handgepäcks und Konditionen zur flexiblen Änderung der Reisedaten;
- Unfreundliches Personal am Check-In und in der Kabine;
- Falsches Spezialmenu oder Fehlverhalten von Mitpassagieren auf der Flugreise;
- Überbuchung oder Flugabsage;
- Probleme bezüglich Upgrade mit gesammelten Meilen in die nächsthöhere Klasse.

Der fehlende Meilen-Upgrade als Szenario, das in der E-Mail-Umfrage ermittelt wurde, wird aufgrund der Expertengespräche bestätigt. Zudem zeigt das Beschwerdeverhalten, dass Passagiere ungenügend über die Tarifbedingungen und -optionen informiert sind.

Dimensionen einer Bestrafung

Basierend auf den Expertengesprächen, dem theoretisch-konzeptionellen Teil dieses Kapitels und dem provisorischen Wirkungsmodell werden die Dimensionen des Objekts *Strafe* im Airline-Bereich in Abbildung 3-4 dargestellt. Mit der dimensional Analyse können die Einzelheiten der Realität bezüglich der Kundenbestrafung herausgearbeitet werden (Schnell et al., 1999). Im Zentrum steht nicht die Sicht des Individuums, sondern das *System Kundensteuerung*.

Ausgangspunkt der dimensional Analyse ist die Motivation der Firma, welche die Kundensteuerung eingeführt hat. Nachdem eine Richtlinie oder ein Anreiz kommuniziert und vom Kunden erlebt wurde, nimmt er sie wahr und reagiert dementsprechend (exogener Einfluss). Die Fairness-Wahrnehmung wird dabei von kundenspezifischen Dispositionen beeinflusst (endogener Einfluss).

Abbildung 3-4: Dimensionsanalyse zur Kundensteuerung

Legende: DL = Dienstleistung

Quelle: Eigene Darstellung

3.5 Beobachtung am Check-In

Um neue Zusammenhänge bezüglich der Kundenreaktion auf negativ wahrgenommene Ereignisse (resp. Strafen) zu entdecken, wird eine deduktive, wissenschaftliche Beobachtung durchgeführt. Sie ist zielgerichtet, systematisch und methodisch kontrolliert (vgl. Greve und Wentura, 1997; Huber, 1995). Die Voraussetzung dazu, eine Theorie oder eine Hypothese, ist mit dem provisorischen Wirkungsmodell gegeben (vgl. Kap. 2.6). Gemäss der Klassifikation von Schnell et al. (1999) und Greve und Wentura (1997) handelt es sich in dieser Untersuchung um eine *verdeckte, nicht teilnehmende, natürliche* und *systematische* Beobachtung (vgl. A.2.1). Die *Systematik* betrifft die Beobachtungskategorien. Ein Kategoriensystem weist folgende drei formalen Anforderungen auf (Kromrey, 2006): (1) Eine Kategorie darf sich nur auf eine Bedeutungsdimension beziehen (Vergleichbarkeit); (2) Kategorien und Unterkategorien müssen sich ausschliessen (Klassifizierbarkeit) und (3) die Kategorien müssen erschöpfend und unabhängig voneinander sein (Vollständigkeit).

Bei der Beobachtung existiert im Gegensatz zu anderen Datenerhebungsmethoden keine Population (Stier, 1999). Die Population der zu beobachtenden Ereignisse konkretisiert sich erst im Laufe der Beobachtung. Das heisst, die Stichprobengrösse oder Kriterien zur Stichprobenwahl können nicht *ex ante* definiert werden. Ein Auswahlplan einer Beobachtung kann sich nur auf Zeiteinheiten und räumliche Einheiten beziehen. Mit dieser Auswahl wird implizit angenommen, dass damit eine Zufallsauswahl bezüglich der eigentlich interessierenden Ergebnisse erfolgt (Kromrey, 2006). Vor einer Beobachtung sind daher die Beobachtungseinheiten (Kategorien) und der Beobachtungszeitraum sowie die räumliche Abgrenzung vorzunehmen. Dabei ist zu beachten, dass bestimmte, aus-

sergewöhnliche Ereignisse die Ergebnisse einer Beobachtung verfälschen können. Daher besteht bei Stichprobenentnahme für eine Beobachtung stets ein unberechenbares Risiko. Dieses Risiko kann nicht durch die Vergrößerung der Stichprobe ausgeglichen werden (Schnell et al., 1999). Folglich kann auch das Ergebnis einer Beobachtung nicht generalisiert werden (Greve und Wentura, 1997).

Im Gegensatz zu Datenerhebungsmethoden wie der Inhaltsanalyse oder der Befragung weist die Beobachtung generell über eine tiefere Validität und Reliabilität auf (vgl. A.2.1). Dennoch können mit deduktiver Beobachtung neue Hypothesen / Erkenntnisse über einen Sachverhalt gewonnen werden, die mit anderen Datenerhebungsmethoden nicht möglich sind. Foster und Cone (1980, S. 332) bestätigen: *“Despite its problems, direct observation is perhaps the most valuable assessment tool of behavioral researchers.”*

3.5.1 Erkenntnis- und Forschungsobjekt

Zur Validierung der Ergebnisse der ersten beiden Schritte (E-Mail-Umfrage und Leitfadeninterview) sowie zur Entdeckung neuer Zusammenhänge drängt sich eine Beobachtung der *Passagiere am Check-In* auf. Mit der Beobachtung wird direkt am Forschungsobjekt „Passagier“ geforscht, ohne eine künstliche Situation zu erzeugen (Kundenperspektive). Zur Untersuchung von emotionalen Zuständen und dem Kundenverhalten eignet sich die Beobachtung als Forschungsmethode (Machleit und Eroglu, 2000).

3.5.2 Forschungsziel

Mit der Beobachtung wird konkretes Kundenverhalten untersucht und der Check-In als potentielle Untersuchungs-Stufe im Airline-Serviceprozess evaluiert (für die quantitative Empirie). Aus der Perspektive des Forschungsmodells werden die aus dem Expertengespräch evaluierten Variablen *Kommunikation* sowie *Ticketpreis* als unabhängige Variablen überprüft.

3.5.3 Analytisches Raster zur Datenauswertung

Das Kategoriensystem (analytische Raster) für die Beobachtung wird bezüglich der Struktur basierend auf der dimensionalen Analyse entwickelt (vgl. Kap. 3.4.5). In Zusammenarbeit mit den Experten wurden elf Ereignisse als kritisch identifiziert und mit Kategorien abgebildet (vgl. Tab. 3-5, Abschnitt 2). Auf das Ereignis können Mitarbeiter verschieden reagieren, weshalb sechs verschiedene Kategorien eingeführt werden (Abschnitt 3). Der Reaktion der Firma stehen verschiedene Verhaltensreaktionen sowie die emotionale Reaktion der Kunden gegenüber (Abschnitt 4). Es ist schwierig, während einer Beobachtung die emotionale Reaktion von Passagieren zu messen. Aus diesem Grund wird eine einfache 5er Likert-Skala mit je zwei Ausprägungen für negative und positive emotionale Reaktionen eingeführt. Die Skala basiert auf der Emotions-Messskala von Oliver und Swan (1989).

Tabelle 3-5: Kategoriensystem Beobachtung

1. Disposition Passagier		3. Kommunikation und Reaktion Mitarbeiter	
- Airline / Check-In	- Gruppengrösse	- Mangelhafte, unklare Erklärung	- Entgegenkommen (Reduktion „Strafe“)
- Klasse		- Keine Reaktion	- Erheben einer Gebühr
- Reisezweck		- Erläuterung mit Grund	- Weiterleiten
- Strecke			
2. Kritisches Ereignis am Check-In		4. Reaktion Passagier	
- Übergepäck	- Falscher Sitzplatz	Verhalten	
- Zu viel Handgepäck	- Passagier nicht aufgeführt	- Akzeptanz: Gebühr zahlen	- Beschwerde
- Sperrgut	- Überbuchung	- Diskussion mit Check-In-Mitarbeiter	- Gepäck: am Flughafen lassen, kumulieren
- Problem mit Ticket / Reisedokumenten	- Problem mit Anrechnung Meilen		
- Problem mit Upgrade	- Check-In bis Enddestination nicht möglich	Emotionale Reaktion	
- Flug verpasst / Flugabsage		wütend; unzufrieden; neutral; zufrieden; glücklich.	

Quelle: Eigene Darstellung

Jeder Zwischenfall wird bei gemeinsam Reisenden einer Gruppe und nicht einer Person zugeordnet (zum Begriff *Gruppe* vgl. Kap. 2.4.3). Eine Gruppe bucht oft über eine Bezugsperson, trägt aber die Folgen einer „Strafe“ als Gruppe (vgl. Tab. 3-5, Abschnitt 1). Die Protokollierung der Beobachtung wurde gemäss Kategoriensystem auf einem Papierbogen vorgenommen.

3.5.4 Datenerhebung

Am Flughafen Zürich gibt es ca. 165 Check-In-Schalter (o.V., 2006d). Der erste Teil der vorliegenden Beobachtung fand bei Jet Aviation am Flughafen Zürich am 11. April 2007 an sieben verschiedenen Check-In-Schaltern statt (vgl. A.5.1). Es wurde sichergestellt, dass flugdemografische Variablen wie der Reisezweck, die gebuchte Klasse und die geflogene Strecke keinen systematischen Einfluss auf die Ergebnisse haben. Beobachtet wurde hinter dem Check-In-Mitarbeiter, wobei der Schalter nach jeder Stunde gewechselt wurde. Dies erlaubte die unauffällige Beobachtung der Passagier-Abfertigung eines *Ethnic Flight (Economy Class)* nach Istanbul (Onur Air), verschiedene Ferienflüge der *Economy* und *Business Class* (Qatar Airways, Edelweiss Air) sowie verschiedene Kurzstreckenflüge der *Economy*, *Business* und *First Class* (British Airways, Air Berlin). Um einen Einblick während einer Zeit mit hohem Passagieraufkommen zu gewinnen, wurde der zweite Teil der Beobachtung am 20. und 21. Juli 2007 durchgeführt (Beginn Sommerferien). Swissport ermöglichte die Beobachtung vom Check-In in der *Economy* sowie der *Business Class* verschiedener Inter- und Kontinentalflüge von Swiss, Lufthansa, Thai Airways und Japan Airlines (total: neun verschiedene Check-In-Schalter).

3.5.5 Forschungsergebnisse

Kundenansprache und Buchung

Die Ergebnisse der Beobachtung bestätigen, dass Passagieren bei der Kundenansprache (z.B. Werbung) und Buchung unklar kommuniziert wird, welche Limiten bezüglich Anzahl Gepäckstücken und Gewicht pro Stück eingehalten werden müssen. 70% der beobachteten Passagiere mit Übergepäck oder zuviel Handgepäck waren im Unklaren über die Gepäcklimiten und äusserten dies am Check-In (vgl. A.5.1).

Das Thema Gewichtslimite ist für Fluggesellschaften von zunehmender Bedeutung. Aufgrund hoher Preise für Frachtgut können Airlines bei Flügen mit weniger Passagieren oder tieferem Gewicht das Flugzeug zusätzlich mit Fracht beladen und damit höhere Erträge erzielen (Jet Aviation, Expertengespräch). Einzelne Airlines beginnen daher, das Durchschnittsgewicht pro Passagier und Gepäckstück auf verschiedenen Destinationen zu berechnen, um die Ladekapazität der Flugzeuge zu optimieren.

Check-In

Zur Beobachtung der Kundenreaktionen auf verschiedene negative Ereignisse bietet sich der Check-In-Schalter an: In kurzer Zeit kann eine grosse Menge an Passagieren beobachtet werden, es findet eine starke Interaktion zwischen Passagier und Mitarbeiter statt und am Check-In werden wichtige, nachgelagerte Airline-Serviceprozesse initiiert. Halten sich Kunden nicht an die Richtlinien, wird am Reisetag am Check-In „bestraft“.

Von insgesamt 478 beobachteten Check-In gab es bei 42 Gruppen (9%) einen Zwischenfall. Der grösste Teil der Zwischenfälle fiel auf das Ereignis *Übergepäck* (27 Fälle) und *Anzahl Gepäckstücke* (5 Fälle). Im Falle von Übergepäck haben 15 Gruppen die Gebühr bezahlt, 11 Gruppen das Gepäck mit Mitreisenden kumuliert und eine Gruppe liess Gepäck am Flughafen stehen, um der Gebühr zu entgehen. Den höchsten Anteil an Übergepäck oder zuviel Handgepäck haben Privatreisende, insbesondere die Ferienreisenden (18 Fälle) und der *Ethnic Travel* (11 Fälle).

In den meisten Fällen (71%) reagierten die Passagiere neutral oder verständnisvoll auf ein Ereignis am Check-In, knapp ein Viertel reagierte unzufrieden oder sogar verärgert. Die Reaktionen fielen je nach Flugklasse oder Bestimmungsort verschieden stark aus. Zur Entwicklung des zentralen Forschungsmodells wird daher angenommen, dass flugdemografische Variablen, und damit indirekt der Ticketpreis, die Kundenreaktion beeinflussen (vgl. Kap. 4.2).

Angestellte von Abfertigungsgesellschaften kannten in sechs Fällen die Richtlinien der Airlines für Übergepäckstrafungen nicht. Die Richtlinien und Limiten werden regelmässig angepasst und sind je nach Destination und Airline verschieden. In diesen Fällen wurde eine mangelhafte Kommunikation der Richtlinien durch die Airlines und Check-In Mitarbeiter beobachtet, was bei den Passagieren zu Unverständnis führte. Unter „mangelhaft“ werden nicht erläuterte Gepäckbestimmungen bei der

Reservation / Kauf des Tickets, verwirrende Kommunikation der Ausnahmeregelungen oder fehlende Hinweise bei Nachfragen verstanden. Die deutliche und faire Kommunikation der Richtlinien vor dem Check-In wird als Einflussfaktor im Wirkungsmodell bestätigt.

Die Beobachtung hat verdeutlicht, dass negative Verhaltensabsichten nicht unmittelbar nach einem Ereignis folgen. Den Passagieren wird bei einem negativen Ereignis wie Übergepäck die Möglichkeit gelassen, eine Gebühr zu entrichten. Damit wird es möglich, nachträglich eine Richtlinie zu umgehen, indem bspw. an einem separaten Airline-Schalter die Übergepäckgebühr bezahlt wird. Die Neigung, eine Strafgebühr zu zahlen oder eine Anreiz-Option zu kaufen, wird als *Akzeptanz* bezeichnet. Das Akzeptanzniveau der Kunden bildet somit ein weiteres Element im Entscheidungsprozess von Kunden, die mit einer Steuerungsmassnahme konfrontiert werden. Das provisorische Wirkungsmodell wird mit der Verhaltens-Variable *Akzeptanz* erweitert (vgl. Kap. 4.2).

Neben Zwischenfällen bezüglich Gepäck wurden beobachtet: 3 x Sperrgut, 3 x ungültige Airline Tickets, 1 x Flug verpasst, 1 x falscher Sitzplatz (Reservation nicht beachtet), 2 x nicht als Fluggast aufgeführt und 2 x Buchung bis an Enddestination nicht möglich. Die detaillierten Ergebnisse der Beobachtung sind im Anhang A.5.1 aufgeführt.

3.6 Inhaltsanalyse Airline-Buchungsprozess

Die *Inhaltsanalyse* ist eine Form der Datenerhebung und der Analysetechnik (Latz, 1993). Als Technik der Datenauswertung wird nach vorgegebenen Kriterien eine Zerlegung des „Bedeutungsträgers“ in Elemente durchgeführt (Atteslander, 2000). Das Ziel ist, Schlüsse bezüglich des analysierten Dokuments ziehen zu können (Kromrey, 2006). Im Gegensatz zur Beobachtung kann die Inhaltsanalyse zu einem beliebigen Zeitpunkt wiederholt werden (Schnell et al., 1999). Zentral ist bei der Inhaltsanalyse das analytische Raster (Kategoriensystem). Es stellt das Raster dar, gemäss dem der Inhalt des Airline-Buchungsprozesses zerlegt wird, und es gewährleistet eine hohe Reliabilität (Stier, 1999).

3.6.1 Erkenntnis- und Forschungsobjekt

Im letzten Schritt der qualitativen Empirie wird eine Inhaltsanalyse anhand verschiedener *Airline-Online-Buchungsprozesse* als Forschungsobjekt durchgeführt, weil sich das Kundenverhalten über die Online-Buchung steuern lässt (vgl. Kap. 3.1.2). Eine erste (exploratorische) Analyse von Airline-Buchungsprozesse wurde bereits im Jahr 2006 durchgeführt. Sie zeigte im Vergleich zur (vorliegenden) Analyse vom 8. Januar 2008, dass Airlines begonnen haben, den Online-Buchungsprozess bezüglich der Kundensteuerung anzupassen. Airlines wie z.B. die Swiss haben den Fokus stärker auf die Kundensteuerung und damit auf eine deutlichere Kommunikation der Tarifbedingungen und Optionen gerichtet.

3.6.2 Forschungsziel

Mit der Inhaltsanalyse können Durchschnittspreise für die experimentelle Manipulation des *Ticketpreises* sowie Preise für die Szenarien im Fragebogen definiert werden (vgl. Kap. 5.2.2). Neben den Durchschnittspreisen für Tickets wird untersucht, ob und welche Anreiz-Optionen sowie Bestrafungsrichtlinien im Airline-Bereich kommuniziert werden. Dabei soll aufgezeigt werden, wie deutlich die Tarifbedingungen und -optionen über das Internet dargestellt sind. Es werden erste Eindrücke für die optische Gestaltung und *Kommunikation* der Website für das web-gestützte Experiment gewonnen.

3.6.3 Analytisches Raster zur Datenauswertung

Immer mehr Passagiere buchen ihren Flug online, direkt bei der Airline oder über ein Webportal (vgl. Kap. 3.1.2). Bei der Inhaltsanalyse der Online-Buchung wird die Art der Kommunikation von Anreiz-Optionen und Richtlinien untersucht (vgl. Tab. 3-6). Die Ergebnisse der Inhaltsanalyse werden im quantitativen Teil verwendet und sind detailliert im Anhang (A.5.2) aufgeführt.

Tabelle 3-6: Analytisches Raster für die Inhaltsanalyse

Airline und GM	Strecke	Ticketpreis (Economy)	Kommunikation Bedingungen	Preis (Strafe)	Steuerung Anreize	Preis (Option)
Network C., LCC, Charter A., Regional A.	Langstrecke oder längste Strecke ab Heimflughafen.	Hin- und Rückflug, billigster Tarif in CHF ohne Taxen.	Anzahl Klicks bis zur Information, Deutlichkeit und Strukturierung.	Übergepäckgebühr in CHF und pro Kilo.	Ja, nein oder teilweise (Angebot an Anreiz-Optionen).	Beschreibung der Option mit Preis in CHF.

Legende: GM = Geschäftsmodell, C. = Carrier, A. = Airline;

Quelle: Eigene Darstellung

3.6.4 Datenerhebung

Verschiedene Airlines haben im Jahr 2006 und 2007 den Online-Buchungsprozess und teilweise die Kundensteuerung über das Internet angepasst. Aus diesem Grund wurde die web-gestützte Inhaltsanalyse zweimal durchgeführt: im Dezember 2006 (exploratorisch) und im Januar 2008 (definitiv). Die insgesamt zwölf Airlines wurden basierend auf dem *Airline Ranking by Passenger Traffic 2006* des Airline Business Magazins ausgewählt (Sobie, 2007). Für die Stichprobenauswahl wird darauf geachtet, dass verschiedene Geschäftsmodelle und Airlines aus verschiedenen Regionen berücksichtigt werden.

3.6.5 Forschungsergebnisse

Im Gegensatz zur Buchung über Reisebüros oder externe Webportale kann die Airline den eigenen Internet-Auftritt und damit die Kundensteuerung selbst gestalten. Die Analyse der Airline Websites

ergab, dass von den zwölf untersuchten Online-Buchungsseiten nur sechs die Kunden klar und deutlich auf die Anreiz-Optionen und Richtlinien aufmerksam machen (vgl. A.5.2). Ein Test der Airline E-Shops in der Schweiz bestätigt diese Erkenntnis (vgl. Steiner, 2007). Die Produktbeschreibung ist bei fünf von acht getesteten Seiten mangelhaft umgesetzt.

Die in der vorliegenden Inhaltsanalyse ermittelten Durchschnittspreise für einen Flug in der *Economy Class* auf der Langstrecke wird für die Manipulation der Variablen *Ticketpreis* im Fragebogen zur quantitativen Untersuchung verwendet. Der Höchstpreis wird mit CHF 3'000.- und der Tiefstpreis mit CHF 400.- festgesetzt. Diese Preise erlauben die Manipulation des Ticketpreises im Experiment („hoch“ vs. „tief“; vgl. Kap. 5.2.2). Sie basieren auf realitätsnahen Durchschnittspreisen verschiedener Airlines zu verschiedenen Buchungszeitpunkten.

3.7 Empirische Grenzen

Die qualitativ-empirische Untersuchung umfasst bei der E-Mail-Umfrage 31 Personen und bei den Expertengesprächen elf Personen. Für qualitative Aussagen und für Hinweise zum Konstrukt *Strafe* sowie als Vorbereitung zur quantitativen Empirie ist dies ausreichend (vgl. Kap. 1.5.2). Es lassen sich jedoch keine Allgemeinaussagen machen, da bei der Datenerhebung nicht auf die Repräsentativität und Stichprobengrösse geachtet wurde.

Aus Anbietersicht werden nur Airline-Experten von Qualitätsairlines mit dem Geschäftsmodell *Network Carrier* sowie Experten aus flughafennahen Bereichen am Flughafen Zürich befragt. Eine Erweiterung der Interviews mit Experten anderer Airlines (Geschäftsmodellen) aus weiteren Regionen würde die Validität und den Erkenntnisgewinn erhöhen. Durch Interviewereffekte kann das Antwortverhalten zudem beeinflusst worden sein, insbesondere, weil das provisorische Wirkungsmodell aus Gründen der Verständlichkeit vor der Befragung präsentiert wurde.

Aus Kundensicht müssten zur Erhöhung der Reliabilität und Validität Beobachtungen über das ganze Jahr hinweg und an verschiedenen Flughäfen durchgeführt werden. Für den Untersuchungsfokus der vorliegenden Arbeit reichen die Daten aus, da sie hauptsächlich zur Bildung des zentralen Forschungsmodells verwendet werden. Das Forschungsmodell wird nachträglich quantitativ überprüft.

Die Ergebnisse der Inhaltsanalyse von Airline Websites unterliegen zeitlichen Schwankungen, da Airlines ihre Websites erfahrungsgemäss nach wenigen Jahren anpassen. Zudem können Verzerrungen auftreten, da lediglich die Untersuchungsobjekte nach den Kriterien *Geschäftsmodell* und *Region* ausgewählt wurden. Eine Vollerhebung wäre aussagekräftiger, aber mit hohem Aufwand verbunden.

3.8 Folgerung

In diesem Kapitel werden die Forschungsfragen b.1 (Einflussfaktoren im Forschungsmodell) und b.2 (Kundensteuerung in der Airline-Branche) implizit aufgrund der Ergebnisse der qualitativen Empirie beantwortet. Die explizite Beantwortung der Forschungsfragen wird in Kapitel 7.2 vorgenommen. Basierend auf den neu integrierten Variablen werden Hypothesen (h) erarbeitet. Sie sind in Kapitel 4.2.2 inhaltlich-theoretisch formuliert.

Einflussfaktoren im Forschungsmodell

Das Kundenverhalten wird bei Airlines mit dem Beschwerdemanagement und mit Kundenumfragen analysiert. Dabei wird missachtet, dass Kunden, die sich illoyal verhalten oder sich negativ über eine Airline äussern, mit dem Beschwerdemanagement nicht zu erfassen sind. Forschung zur Wiedergutmachung im Airline-Bereich hat gezeigt, dass eine schnelle Entschuldigung und Behebung des Schadens eine effiziente Strategie ist. Zusätzliche Kompensation ist nicht nötig (de Coverly et al., 2002).

Das zentrale Forschungsmodell für die quantitative Untersuchung mit den zugehörigen Hypothesen wird basierend auf Erkenntnissen dieses Kapitels aufgebaut. Die qualitative Untersuchung bestätigt das in Kapitel 2.6 erarbeitete, provisorische Wirkungsmodell. Insbesondere wird die Variable Fairness-Beurteilung als Indikator für die individuelle Wahrnehmung der *Kundensteuerungsmassnahme* bestätigt (h1). Es wird nachgewiesen, dass zur korrekten Wahrnehmung der Kundensteuerungsmassnahme die *Kommunikation* zentral ist (h2).

Das provisorische Wirkungsmodell wird mit der Integration neuer Variablen aus der qualitativen Empirie und aus theoretischen Grundlagen im Airline-Bereich zum *zentralen Forschungsmodell* erweitert (vgl. Abb. 3-5). Die Beziehung zwischen der Kundensteuerung, deren Kommunikation und der Fairness-Beurteilung kann von folgenden Variablen zusätzlich beeinflusst werden: *Ticketpreis* (h3), *Image der Airline* (h4) vor einer Flugreise (*ex ante*), *Airline-Geschäftsmodell* (h8) und *flugdemografische Variablen* (h_d). Da sich nach einem kritischen Ereignis oder einer Bestrafung die Image-Wahrnehmung ändert, wird im Modell die *ex post Image* Variable (h9) als abhängige Variable eingeführt (vgl. Kap. 2.1.1). Ebenfalls berücksichtigt wird die abhängige Variable *Akzeptanz* (h6), welche die Zahlungsbereitschaft für eine Option oder eine Strafe aufzeigt (vgl. Kap. 4.2).

Kundensteuerung in der Airline-Branche

Richtlinien mit Strafcharakter haben sich in der Airline-Branche etabliert. Eine Airline steuert heute das Kundenverhalten vor allem über Push-Massnahmen, die beim Kauf eines Tickets entweder im Reisebüro oder auf der Website als Tarifbedingungen (oft unklar) kommuniziert werden. Die Zunahme von Konkurrenz auf einzelnen Strecken und neue Geschäftsmodelle mit Leistungsvorteilen ermöglichen unzufriedenen Passagieren, sich illoyal zu verhalten. Es sollten vermehrt Anreiz-Optionen, die heute (oft) als Richtlinien formuliert sind, angeboten werden. Sie dienen als Diffe-

Abbildung 3-5: Beitrag der qualitativen zur quantitativen Empirie und zum Forschungsmodell

	Forschungsergebnisse der qualitativen Empirie	Beitrag zu quant. E. / Forschungsmodell
E-Mail-Umfrage	<ul style="list-style-type: none"> – Fairness-Beurteilung als Indikator für Kundensteuerung – Richtlinien als „Strafe“ empfunden, emotionale Reaktion – Szenario: Meilen-Upgrade anwendbar 	<ul style="list-style-type: none"> – Prov. Modell bestätigt – Fragebogenentwicklung: Szenario Meilen-Upgrade
Experten-interview	<ul style="list-style-type: none"> – Buchung über Internet als Kundensteuerung „Anreiz“ – Adaption Hypothesen auf Praxisnähe – Szenario: Lounge / VIP Boarding und flexibles Ticket 	<ul style="list-style-type: none"> – Entwicklung Forschungsmodell und Fragebogen – Hypothesen (h4, h6, h8, h9)
Beobachtung	<ul style="list-style-type: none"> – Check-In als Stufe zur Kundensteuerung „Richtlinie“ – Kommunikation, Ticketpreis und flugdemografische Variablen als Einflussfaktoren – Akzeptanz der Gebühr als Verhalten im Modell – Szenario: Übergepäck anwendbar (Emotionen) 	<ul style="list-style-type: none"> – Entwicklung Forschungsmodell, Hypothesen 1 - 3 und demografische H. – Fragebogenentwicklung: Szenario Übergepäck
Inhalts-analyse	<ul style="list-style-type: none"> – Mangelhafte Kommunikation Optionen / Bedingungen – Durchschnittspreise für Ticket und Option 	<ul style="list-style-type: none"> – Fragebogen: Manipulation der UV, Preise der Optionen in Szenarien

Legende: H. = Hypothesen, quant. E. = quantitative Empirie, Prov. = provisorisches;

Quelle: Eigene Darstellung

renzierungsmerkmale gegenüber dem Airline-Branchen-Standard. Dabei steht nicht das Bonusprogramm im Vordergrund, sondern Zusatzoptionen zum Flugticket. Zur quantitativen Untersuchung der Kundensteuerungsmassnahmen eignen sich aus Sicht der Airline-Servicekette die Stufen *Reservation / Kauf Ticket im Internet* für die Anreiz-(Pull-) Steuerung und *Check-In* für die Steuerung über Richtlinien (Push-Massnahme). Vier verschiedene Szenarien werden erarbeitet, die von der Airline sowohl als Anreiz-Optionen als auch als Bestrafungsrichtlinien gestaltet werden können. Sie werden experimentell mit dem quantitativen Teil überprüft:

- *Übergepäck* als Option über das Internet anbieten oder am Check-In mit Gebühr steuern;
- Anrecht auf *Lounge-Zugang* sowie *VIP Boarding* (Boarding mit Priorität) über Internet anbieten oder am Check-In auf fehlendes Anrecht hinweisen (Tarifbedingungen);
- *Flexibles Rückreisedatum* als Internet-Option anbieten oder am Check-In Restriktion auf Umbuchungsmöglichkeiten mitteilen (Tarifbedingungen);
- *Upgrade* in die nächsthöhere Klasse mit Meilen aus dem Vielfliegerprogramm über Internet anbieten oder am Check-In *Upgrade* nicht durchführen (z.B. aus organisatorischen Gründen).

4 Forschungsmodell und Untersuchungsfokus

In Kapitel 4.1 werden die Grundlagen der Modell- und Hypothesenbildung erarbeitet. Für die vorliegende quantitative Forschung werden in Kapitel 4.2 inhaltlich-theoretische (Haupt-) Hypothesen und Nebenhypothesen formuliert. Die Nebenhypothesen betreffen demografische Variablen und Interaktionseffekte, die nicht experimentell manipuliert werden. Die Hypothesen geben (vorläufige) A-priori-Antworten auf die Forschungsfragen und basieren auf den theoretisch-konzeptionellen Grundlagen und dem provisorischen Wirkungsmodell. Das provisorische Wirkungsmodell wurde mit dem qualitativen Forschungsteil überprüft und für den quantitativen Teil bezüglich der Hypothesenformulierung erweitert (im Sinne eines relativistisch-konstruktivistischen Wissenschaftsverständnisses). Der Untersuchungsfokus und damit die Methodenauswahl werden in Kapitel 4.3 diskutiert. Die quantitative Untersuchung wird aufgrund der Kausalbeziehungen im Forschungsmodell mit einem experimentellen Design durchgeführt (vgl. Kap. 4.3.2).

4.1 Grundzüge der Modell- und Hypothesenbildung

Beim vorliegenden Forschungsmodell (vgl. Abb. 4-1) handelt es sich um ein *Kausalmodell*. In Kausalmodellen werden Variablen und Konstrukte sowie deren gegenseitige Beziehung dargestellt, wobei bestimmte Variablen als ursächlich für die Ausprägung anderer Variablen angesehen werden (Schnell et al., 1999; Bollen, 1989). *Kausalität und Korrelationen* im Modell determinieren die Wirkbeziehung zwischen den Variablen und sind in Kapitel 4.1.1 erläutert.

Im Forschungsbereich des Kundenverhaltens werden *Variablen* als veränderliche Beobachtungsgrößen aus dem Bereich menschlichen Erlebens, Verhaltens und Handelns definiert (Trommsdorff, 2003). Da mehrere abhängige und unabhängige Variablen gleichzeitig im Kausalmodell eingesetzt werden, handelt es sich um ein Kausalmodell *multivariater* Art (Kromrey, 2006). Eine wichtige Rolle nehmen im vorliegenden Forschungsmodell die *Moderatorvariablen* ein (Kap. 4.1.2). Die Ursache-Wirkung-Zusammenhänge zwischen den Variablen werden als *Hypothesen (h)* formuliert (Kap. 4.1.3). Unter einem *Konstrukt* werden theoretische Eigenschaftsdimensionen, resp. *latente* Variablen, verstanden (Stier, 1999). Die latenten Variablen sind nicht direkt, sondern indirekt über Indikatoren messbar (z.B. „Angst“, „Intelligenz“ etc.).

4.1.1 Kausalität und Korrelation

In einem Experiment werden verschiedene Typen von Variablen unterschieden: abhängige und unabhängige Variablen. Die abhängige Variable (AV) ist die Merkmalsdimension, deren Ausprägungsgrad durch das Wirkungsmodell erklärt werden soll (Diekmann, 1998). Das Merkmal, auf dessen unterschiedliche Faktorstufe die Variation der AV zurückzuführen ist, wird unabhängige Variable (UV) oder Faktor genannt (Hirsig, 2001). Die „Wirkungsrichtung“ geht von den unabhängigen zu

den abhängigen Variablen. Es wird auch von *exogenen* (unabhängigen) und *endogenen* (abhängigen) Variablen gesprochen (Kromrey, 2006).

Zwei Variablen sind entweder miteinander korreliert oder kausal verknüpft (Hirsig, 2001). *Kausalität* liegt vor, wenn eine unabhängige Variable ursächlich mit einer abhängigen Variablen zusammenhängt (Byrne, 2001). Die UV ist nicht kausal von anderen Modellvariablen abhängig. Sie kann für die Veränderungen der AV verantwortlich sein (Baron und Kenny, 1986). Eine *Korrelation* hingegen besagt, dass zwei Variablen eine gegenseitige Beziehung aufweisen und sich miteinander verändern (Bortz, 2006). Um Kausalität von Korrelation unterscheiden zu können, ist es wichtig, dass das Modell theoretisch und empirisch untermauert ist. Kausalität kann nicht statistisch überprüft werden (Hirsig, 2001; Bollen, 1989). Daher werden die Kausalbeziehungen des vorliegenden Modells einerseits mit dem theoretischen Bezugsrahmen und andererseits mit der qualitativ-exploratorischen Forschung begründet.

4.1.2 Moderatorvariablen

Eine *Moderatorvariable* verändert den Zusammenhang zwischen der unabhängigen und der abhängigen Variable. Der Moderator ist eine Drittvariable, die die Stärke einer Wirkbeziehung zwischen zwei Variablen beeinflusst (Baron und Kenny, 1986). Die Richtung und Intensität der Korrelation zwischen UV und AV werden beeinflusst. Im Gegensatz zum Moderator „vermittelt“ der *Mediator* indirekt zwischen UV und AV und hängt bedeutsam mit beiden Faktoren zusammen (James und Brett, 1984).

4.1.3 Experimentelle Hypothesen

Eine Hypothese ist eine in spezielle Aussageform gekleidete Fragestellung (Schnell et al., 1999). Sie ist eine präzisierte Vermutung mit Vorhersagen zu „Wenn-Dann“-Beziehungen (Czienskowski, 1996). Über deren Gültigkeit oder Ungültigkeit kann der Experimentator aufgrund seiner Befunde entscheiden. Dabei werden die in der Hypothese formulierten Aussagen mit den unter kontrollierten Bedingungen gesammelten Beobachtungsdaten konfrontiert (Hussy und Jain, 2002). Für die Qualität der Hypothesen spricht, wenn diese den kritischen Prüfungen standhalten. Hypothesen sind im Rahmen des hypothetisch-deduktiven Vorgehens das Bindeglied zwischen den Theorien und den empirischen Daten.

Da Hypothesen auch unsystematisch „entdeckt“ werden können (kreativer Zufall), entstand die bereits erläuterte Kritik Feyerabends am systematischen, theoriegeleiteten Vorgehen (vgl. Kap. 1.5.1). Trotzdem bedeutet diese Tatsache im konkreten Fall, dass eine „systematische“ Beherrschung der Theorie und der Methode nötig ist (Winer et al., 1991). Empirisch sinnvolle Hypothesen müssen immer an der Erfahrungswirklichkeit testbar sein. Die Möglichkeiten der experimentellen Methode sind besonders eng an den Einfallsreichtum des Experimentators geknüpft (Maxwell und Delaney,

1990). Es gilt, Fragestellungen derart in Hypothesenform zu formulieren, dass das Experiment eine Antwort auf die Fragestellung geben kann (Hicks, 1993).

Bei einem experimentellen Design besteht ein enger Zusammenhang zwischen der Formulierung der Hypothesen und der Manipulation konkreter unabhängiger Variablen sowie der Beobachtung und Messung von einzelnen abhängigen Variablen. Die Grenzen der Manipulation in einem Experiment bestimmen daher die Formulierung empirisch überprüfbarer Hypothesen (Czienskowski, 1996). Folgende Gütekriterien müssen für wissenschaftlich relevante Hypothesen erfüllt sein:

- *Theorierelevanz*: Aus Theorien oder Modellen abgeleitete Hypothesen erhöhen den „Beziehungsreichtum“ einzelner Hypothesenbildungen im Vergleich zu Hypothesen, die nicht von Theorien oder Modellen abgeleitet sind (Czienskowski, 1996; Kirk, 1995).
- *Testbarkeit und Falsifizierbarkeit*: Je präziser und informationsreicher Hypothesen formuliert sind, desto günstiger ist die Voraussetzung zur Falsifikation (Czienskowski, 1996).
- *Präzisierung*: Mathematisch formulierte Null- und Alternativhypothesen sind leichter falsifizierbar als globale Hypothesen. Die Nullhypothese postuliert, dass beide Stichproben aus derselben Population stammen. Die Alternativhypothese ist das logische Komplement zur Nullhypothese, lässt sich aber nicht statistisch bearbeiten (vgl. A.3.2).
- *Occam's Razor (Prinzip der Sparsamkeit)*: Bei sonst gleichem Voraussagewert sollen einfachere, sparsamere Grundlagen solchen vorgezogen werden, die eine kompliziertere und damit auch spekulativere Basis der Hypothesenbildung zur Voraussetzung haben (Winer et al., 1991).

In der vorliegenden Arbeit haben die Hypothesen ihren Ursprung in der Auseinandersetzung mit dem theoretisch-konzeptionellen und dem qualitativ-exploratorischen Teil; sie dienen der Verallgemeinerung von bisher bekannten Einzelfakten und zur Bestätigung des Forschungsmodells (deduktives Vorgehen). Aufgrund der Forschungsperspektive wird in dieser Arbeit die Wirkungsrichtung (größer, kleiner) der Hypothesen definiert (Latz, 1993). Die Hypothesen werden empirisch überprüft und sind im Modell als Pfeile dargestellt (vgl. Abb. 4-1).

Die Operationalisierung der Variablen erfolgt unter Bezugnahmen von Hintergrundwissen (vgl. Kap. 5.2). Dieses Wissen enthält die zur Zeit in einem spezifischen Wissenschaftsbereich bekannten Gesetzmässigkeiten, empirischen Resultate oder Vorannahmen (Hussy und Jain, 2002). Eine Hypothese, bei welcher noch keine Operationalisierung der abstrakten Variablen vorgenommen wurde, wird theoretisch-inhaltliche Hypothese genannt. Durch Einbezug von Hintergrundwissen wird sie zu einer empirisch-inhaltlichen Hypothese und statistisch bearbeitbar (Kirk, 1995).

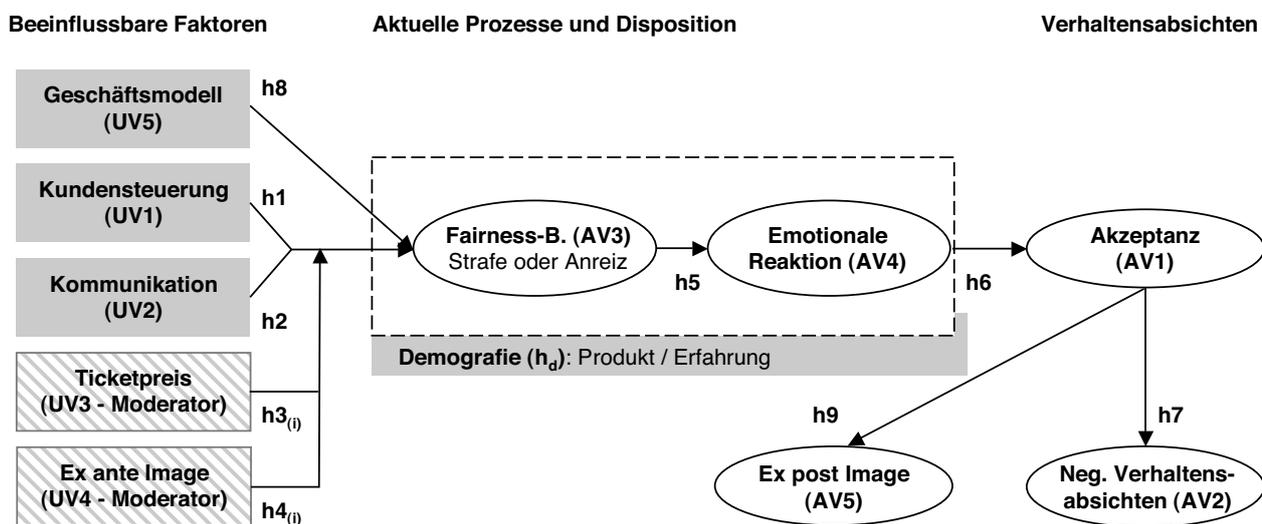
Im nächsten Kapitel werden die Hypothesen zum vorliegenden Forschungsmodell diskutiert. Die Operationalisierung der Variablen wird in Kapitel 5.2.2 und 5.2.3, die statistisch bearbeitbaren Testhypothesen sind in Kapitel 6.3 und 6.5 für beide Experimente erläutert.

4.2 Forschungsmodell und Hypothesen

Das zu testende, zentrale Forschungsmodell basiert auf dem provisorischen Wirkungsmodell (vgl. Kap. 2.6). Das provisorische Wirkungsmodell wird aufgrund von Erkenntnissen der qualitativen Empirie und Forschung im Airline-Bereich (vgl. Kap. 3.8) um die Faktoren *Ticketpreis*, *ex ante Image*, *Geschäftsmodell* und die Variablen *Akzeptanz* sowie *ex post Image* erweitert.

Die Grundlagen zur Variablenselektion sind in Kapitel 4.2.1 aufgeführt. Im zentralen Forschungsmodell werden Faktoren untersucht, die im Einflussbereich der Airline sind (*beeinflussbare Faktoren*) und deren Auswirkungen auf die *aktuellen Prozesse* sowie *Verhaltensabsichten* von Passagieren (vgl. Abb. 4-1). Die Wirkungszusammenhänge werden als inhaltlich-theoretische Hypothesen in Kapitel 4.2.2 modelliert.

Abbildung 4-1: Zentrales A-priori-Forschungsmodell



Legende: AV = abhängige Variable, UV = unabhängige Variable, B. = Beurteilung, h = Hypothese;
Quelle: Eigene Darstellung

Das Forschungsmodell wird mit zwei Experimenten getestet, um die Güte der Untersuchung zu erhöhen (vgl. Kap. 4.3.2). In einem ersten Schritt werden Probanden direkt am Flughafen mit einem Fragebogen befragt, wobei die Hypothesen 1 bis 7 überprüft werden (vgl. Kap. 4.2.2). Mit dem web-gestützten Experiment werden in einem zweiten Schritt die Hypothesen 8 bis 10 getestet (für Hypothese 10 vgl. Kap. 6.4.3). Zusätzlich werden mit dem zweiten Experiment die zentralen Hypothesen 1, 2 und 7 validiert. Die (flug-) demografischen Variablen (d = Disposition) sowie allfällige Interaktionseffekte (i) zwischen den unabhängigen Variablen werden nicht experimentell untersucht. Als Nebenhypothesen (h_i und h_d) modelliert, kann ihr Einfluss dennoch überprüft werden (vgl. Kap. 4.2.2). Die demografischen Variablen werden im Modell nicht als unabhängige Variablen abgebildet, da es sich um endogene Variablen handelt, die nicht experimentell untersucht werden.

4.2.1 Variablenselektion

Im Bereich des Kundenverhaltens und der Kundenwahrnehmung bestehen mehrere Theorien und Variablen zur Erklärung des gleichen Phänomens (vgl. Kap. 2.1.4; Kroeber-Riel und Weinberg, 2003). Die Auswahl der untersuchungsrelevanten Variablen für das Forschungsmodell stellt somit eine Annahme dar. Sie ist eine entscheidende Weichenstellung für den weiteren Forschungsprozess (Kromrey, 2006). In der vorliegenden Untersuchung werden die Variablen in das Forschungsmodell aufgenommen, wenn (vgl. Bortz, 2006):

- (a) die theoretisch-konzeptionellen Grundlagen die Bedeutung der Variable gemäss Forschungsfrage nahelegen und / oder
- (b) die qualitativ-exploratorische Empirie die Bedeutung der Variable gemäss Forschungsfrage bestätigt.

Demnach bestimmt die Forschungsperspektive den Einbezug der Variablen, wobei im besten Fall (a) und (b) zutreffen. Im vorliegenden Forschungsmodell werden fünf unabhängige und fünf abhängige Variablen untersucht. Die ausgewählten Variablen, deren Bedeutung und die Grundlage der Selektion für das quantitativ zu testende Forschungsmodell sind in Tabelle 4-1 aufgeführt.

Tabelle 4-1: Bedeutung und Selektion der Variablen für das Forschungsmodell

Variable	Bedeutung	Selektion
UV1: Kundensteuerung	Massnahme der Airline: Anreiz oder Richtlinie	a (2.1.2; 3.1.3), b
UV2: Kommunikation	Interaktionale Gerechtigkeit der Kundensteuerung	a (2.3.2; 2.3.4), b
UV3: Ticketpreis	Preis einer Flugreise (Langstrecke, Economy Class)	a (2.3.2), b
UV4: Ex ante Image	Marken-Wahrnehmung des Kunden vor Flugreise	a (2.1.1)
UV5: Geschäftsmodell	Beschreibung, wie Airline Werte schafft	a (3.1.1), b
AV1: Akzeptanz	Kundenverhalten bezüglich Zahlungsbereitschaft	b
AV2: Neg. Verhaltensabsichten	Illoyalität, negative WOM und Beschwerde	a (2.1.4)
AV3: Fairness-Beurteilung	Kognitive Wahrnehmung der Kundensteuerung	a (2.3.2), b
AV4: Emotionale Reaktion	Indikator für die Verhaltensabsichten des Kunden	a (2.3.1; 2.5.3), b
AV5: Ex post Image	Marken-Wahrnehmung des Kunden nach Flugreise	a (2.1.1)

Legende: a = theoretischer Bezugsrahmen (Kapitel), b = qualitative Empirie;
Quelle: Eigene Darstellung

4.2.2 Inhaltlich-theoretische Hypothesen

Die Hypothesen werden basierend auf dem provisorischen Wirkungsmodell und Erkenntnissen aus Kapitel 3 aus Anbieter- und aus Kundenperspektive modelliert.

Faktoren im Einflussbereich des Anbieters

Die Hypothesen in diesem Kapitel betreffen fünf unabhängige Variablen (UV1 - UV5), die aus Sicht des Anbieters *manipulierbar* sind, um die kognitive Fairness-Beurteilung beim Individuum zu beeinflussen (vgl. Kap. 2.3.2).

Die zentralen Elemente in dieser Forschungsarbeit sind gemäss Forschungsfragen (vgl. Kap. 1.4.2) das Konzept zur *Kundensteuerung* (h1), die *Kommunikation* (h2) der Kundensteuerungsmassnahme und die Auswirkung auf die *Verhaltensabsichten* (h7). Die Kundensteuerung kann als Anreiz-Option (Pull-Massnahme) bei der Buchung eines Fluges oder als Richtlinie mit Strafe (Push-Massnahme) am Check-In kommuniziert werden (vgl. Kap. 2.1.2; 3.1.4). Die Konsequenz einer nicht befolgten Richtlinie ist die Bestrafung. Sie wird gemäss Theorien zur kognitiven Evaluation, der *Justice, Fairness* und *Prospect Theory* von Kunden als unfair eingestuft (vgl. Kap. 2.3.2; 2.3.4). Daraus leitet sich die erste Hypothese (h) ab:

h1: Wenn die Kundensteuerung als Push-Massnahme gestaltet wird, dann wird aus Kundensicht die Sanktion als Bestrafung (tiefe Fairness-Beurteilung) wahrgenommen.

Wenn Kunden die Ursache für ein Ereignis auf sich selbst attribuieren (Lokus der Kausalität), werden sie mit einer höheren Fairness-Beurteilung und somit mit einer tieferen emotionalen Intensität reagieren (vgl. Kap. 2.3.3). Hierzu muss genügend Vorwissen zur Kundensteuerungsmassnahme bestehen. Mit richtiger Kommunikation wird beim Kunden Vorwissen aufgebaut (vgl. Kap. 3.1.4). Eine deutlich kommunizierte Push- oder Pull-Massnahme führt gemäss Forschungsergebnissen zur Wiedergutmachung und gemäss der interaktionalen Gerechtigkeit zu einer höheren Fairness-Beurteilung (vgl. Kap. 2.1.3; 2.3.2; 3.1.7).

h2: Je deutlicher die Kommunikation, umso höher ist die kognitive Fairness-Beurteilung (Wahrnehmung als Anreiz).

Im Zusammenhang mit den Hauptvariablen Kundensteuerung, deren Kommunikation und der daraus abgeleiteten Fairness-Beurteilung werden in dieser Arbeit zwei moderierende Variablen identifiziert: der *Ticketpreis* (UV3) und das *ex ante Image* der Airline (UV4). Diese Variablen haben im Kontext der Forschungsfragen keine ursächliche Auswirkung auf die kognitive Fairness-Beurteilung, sondern beeinflussen die Wirkbeziehung zwischen der Kundensteuerung, deren Kommunikation und der Fairness-Beurteilung.

Zur Untersuchung des moderierenden Effekts muss zwischen dem Moderator und der abhängigen Variablen ein signifikanter Zusammenhang bestehen (Baron und Kenny, 1986). Daher wird in einem ersten Schritt die Hypothese bezüglich der AV *Fairness-Beurteilung* modelliert (vgl. h3 und h4). In einem zweiten Schritt werden die Interaktionseffekte untersucht (vgl. A.3.3).

Für die Equity-Wahrnehmung spielt der Ticketpreis bei einer Bestrafung eine zentrale Rolle (vgl. Kap. 2.3.2). Im Gegensatz zu *Low-Involvement*-Produkten beschäftigen sich Kunden bei einer teuren Langstreckenreise (*high involvement*) intensiver mit den Tarifbedingungen und Optionen (vgl. Kap. 3.1.5). Von Kunden im höheren Preissegment wird zudem angenommen, dass sie durchschnittlich über eine höhere Erfahrung bezüglich Flugreisen verfügen. Sie können besser abschätzen, ob das Ereignis einmalig ist, nochmals vorkommen wird oder ob es sich um eine branchenübliche Massnahme handelt (vgl. Kap. 2.3.3; 3.1.4). Aufgrund des höheren Vorwissens und der Flugerfahrung folgt eine höhere Fairness-Beurteilung bei einer Bestrafung. Da der Ticketpreis und das *ex ante* Image als Moderatorvariablen identifiziert wurden, werden die dazugehörigen Interaktions-Hypothesen ebenfalls aufgeführt (vgl. auch h4).

h3: Je höher der Ticketpreis, umso höher manifestiert sich die kognitive Fairness-Beurteilung (einer Push- oder Pull-Massnahme).

h3_{i1}: Der Ticketpreis hat einen moderierenden Effekt auf die Kundensteuerung.

h3_{i2}: Der Ticketpreis hat einen moderierenden Effekt auf die Kommunikation.

Image ist ein multidimensionales Konstrukt (vgl. Kap. 2.1.1). Es wird vom Kunden vor dem Kauf einer Flugreise mit einer Airline assoziiert. Ist dieses *ex ante Image* der Airline hoch, wird ein Passagier bei einem Zwischenfall mit Nachsicht reagieren, weil er diesen als „Ausnahme“ taxiert (vgl. Kap. 3.1.7). Verfügt die Airline über ein tiefes *ex ante Image*, wird der Kunde bei einem unerwarteten, kritischen Ereignis mit einer tiefen Fairness-Beurteilung reagieren (vgl. Kap. 2.1.1).

h4: Je höher das *ex ante Image* einer Airline vor dem Ereignis, umso höher manifestiert sich die kognitive Fairness-Beurteilung (einer Push- oder Pull-Massnahme).

h4_{i1}: Das *ex ante Image* hat einen moderierenden Effekt auf die Kundensteuerung.

h4_{i2}: Das *ex ante Image* hat einen moderierenden Effekt auf die Kommunikation.

Mit Hypothese 8 wird postuliert, dass je nach wahrgenommenem Geschäftsmodell (*LCC*, *Network Airline*, *Charter Airline* oder *Regional Airline*) eine Push- oder Pull-Massnahme bezüglich des Fairness-Ratings anders beurteilt wird.⁷ Der moderierende Effekt vom Ticketpreis und vom Image der Airline werden in der Wirkbeziehung zwischen dem Geschäftsmodell und dem Fairness-Rating nicht untersucht. Aus Sicht der Kundenwahrnehmung wird das Airline-Geschäftsmodell in den meisten Fällen mit den Variablen *Ticketpreis* und *Image* kausal verknüpft (vgl. Kap. 3.1.1).

⁷ Hypothese 8 ist nach h4 nicht fortlaufend nummeriert, da sie in der empirischen Untersuchung mit dem zweiten Experiment getestet wird. Die Hypothesen für das erste Experiment sind fortlaufend nummeriert.

Das Geschäftsmodell *Low Cost Carrier* bspw. zeichnet sich durch das Image einer günstigen Airline ohne zusätzlichen Kundenservice (*no frills*) aus.

h8: Wenn Passagiere mit einem anderen Geschäftsmodell konfrontiert werden, dann ändert sich die kognitive Fairness-Beurteilung (einer Push- oder Pull-Massnahme).

Aktuelle Prozesse und Verhaltensabsichten von Passagieren

In der vorliegenden Forschungsarbeit wird davon ausgegangen, dass Individuen vor der emotionalen Reaktion eine kognitive (Fairness-) Beurteilung der Situation vornehmen (vgl. Kap. 2.3.1). Aus den Erkenntnissen zu affektiven Reaktionen, speziell der Frustration (vgl. Kap. 2.5.2) lässt sich somit Hypothese 5 aufstellen:

h5: Je tiefer die kognitive Fairness-Beurteilung, umso stärker manifestiert sich die negative emotionale Reaktion.

Sobald einem Kunden beim Konsum einer Dienstleistung negative Konsequenzen verursacht oder positive Konsequenzen verunmöglicht werden, reagiert er negativ emotional (vgl. Kap. 2.5.3). Zwischen diesen negativen Emotionen und den Verhaltensabsichten besteht ein direkter Zusammenhang (vgl. Kap. 2.4.1). Im Kontext der vorliegenden Untersuchung äussert sich das Verhalten in der Akzeptanz (AV1) einer Anreiz-Option oder einer Strafe: Wird die Option gekauft oder die Strafgebühr bezahlt? Mit Hypothese 6 wird eine gleichsinnige Beziehung zwischen der Emotion und der Akzeptanz postuliert:

h6: Je negativer die emotionale Reaktion, umso tiefer manifestiert sich die Akzeptanz des Kundensteuerungskonzepts.

Die ökonomische Betrachtung zur Bestrafung hat gezeigt, dass es *over-punishing* gibt (vgl. Kap. 2.4.1). Damit geht auch die höhere Wahrscheinlichkeit für negative Verhaltensabsichten einher (vgl. Kap. 2.5.3). Basierend auf dem Basismodell zur Wahrnehmung und dem Equity-Prinzip treten negative Verhaltensabsichten auf, wenn die Massnahme zur Kundensteuerung nicht akzeptiert wird (vgl. Kap. 2.3.2). Daraus folgt die gegensinnige (negative) Beziehung:

h7: Je tiefer die Akzeptanz des Kundensteuerungskonzepts bei Push-Massnahmen, umso höher ist die Wahrscheinlichkeit negativer Verhaltensabsichten.

Die Wahrnehmung des Airline-Images wird *ex post* durch das Kundensteuerungskonzept beeinflusst (vgl. Kap. 2.1.1). Im Forschungsmodell wird dieser Einfluss als Zusammenhang zwischen dem Akzeptanzniveau und der *ex post* Image-Wahrnehmung postuliert:

h9: Je tiefer die Akzeptanz des Kundensteuerungskonzepts, umso tiefer ist die Wahrnehmung des *ex post* Airline-Images.

Einfluss endogener (flug-) demografischer Variablen

Der Einfluss des Passagierverhaltens als demografische Variable wird unabhängig von den experimentellen Bedingungen erhoben und als Demografie-Nebenhypothesen (h_d) geprüft. Im Kontext der forschungsleitenden Fragestellung sind vor allem die Einflüsse von Variablen des Flugverhaltens auf die abhängigen Variablen *Akzeptanz (-niveau)* und *negative Verhaltensabsichten* von Interesse (vgl. Kap. 3.8). Untersucht wird der Einfluss der Klassenzugehörigkeit (*Economy, Business* und *First Class*), Anzahl Flüge im Jahr 2006 (Geschäfts- oder Privatreisen) und Status des Passagiers im Vielfliegerprogramm (vgl. Kap. 3.1.4). Für die Bildung der Nebenhypothesen wird auf dasselbe Hintergrundwissen wie bei h_3 bis h_5 zurückgegriffen (vgl. oben):

- h_{d1} :** Je tiefer die Klassenzugehörigkeit (Produkt), umso tiefer ist die Akzeptanz des Kundensteuerungskonzepts (a) und umso höher ist die Wahrscheinlichkeit für negative Verhaltensabsichten (b).
- h_{d2} :** Je grösser die Flugerfahrung, umso grösser ist die Akzeptanz des Kundensteuerungskonzepts (a) und umso tiefer ist die Wahrscheinlichkeit für negative Verhaltensabsichten (b).

In Tabelle 4-2 sind die Hypothesen und die involvierten Variablen zusammengefasst. Eine Übersicht zu den Nebenhypothesen ist im Anhang aufgeführt (vgl. A.2.3).

Tabelle 4-2: Hypothesen und involvierte Variablen

Hypothesen (h)		Variablen
h1	Wenn die Kundensteuerung als Push-Massnahme gestaltet wird, dann wird aus Kundensicht die Sanktion als Bestrafung wahrgenommen.	UV1 → AV3
h2	Je deutlicher die Kommunikation, umso höher ist die kognitive Fairness-Beurteilung.	UV2 → AV3
h3	Je höher der Ticketpreis, umso höher manifestiert sich die kognitive Fairness-Beurteilung (einer Push- oder Pull-Massnahme).	UV3 → AV3
h4	Je höher das ex ante Image der Airline vor dem Ereignis, umso höher manifestiert sich die kognitive Fairness-Beurteilung (einer Push- oder Pull-Massnahme).	UV4 → AV3
h5	Je tiefer die kognitive Fairness-Beurteilung, umso stärker manifestiert sich die negative emotionale Reaktion.	AV3 → AV4
h6	Je negativer die emotionale Reaktion, umso tiefer manifestiert sich die Akzeptanz des Kundensteuerungskonzepts.	AV4 → AV1
h7	Je tiefer die Akzeptanz des Kundensteuerungskonzepts bei Push-Massnahmen, umso höher ist die Wahrscheinlichkeit negativer Verhaltensabsichten.	AV1 → AV2
h8	Wenn Passagiere mit einem anderen Geschäftsmodell konfrontiert werden, dann ändert sich die kognitive Fairness-Beurteilung.	UV5 → AV3
h9	Je tiefer die Akzeptanz des Kundensteuerungskonzepts, umso tiefer ist die Wahrnehmung des ex post Airline-Images.	AV1 → AV5

Quelle: Eigene Darstellung

4.3 Untersuchungsfokus der quantitativen Empirie

Mit dem Untersuchungsfokus wird der *Untersuchungskontext* der quantitativen Untersuchung erläutert (Kap. 4.3.1), die *Methodenwahl* begründet (Kap. 4.3.2) und die *Forschungsperspektive* mit dem Erkenntnis- und Forschungsobjekt abgegrenzt (Kap. 4.3.3). Das Forschungsziel wird für beide Experimente (Flughafen- und Online-Experiment) einzeln in Kapitel 4.3.4 vorgestellt.

Die Forschungsergebnisse werden in Kapitel 6 präsentiert, nachdem die empirischen Grundlagen für die quantitative Untersuchung erarbeitet wurden (Kap. 5).

4.3.1 Untersuchungskontext

Die qualitative Empirie diente der Exploration des Bestrafungs-Konstrukts, zur Variablenselektion und der Überprüfung der Hypothesen auf Realitätsnähe. Im Gegensatz dazu werden mit der quantitativen Empirie das zentrale A-priori-Forschungsmodell und die dazugehörigen Hypothesen aus Kapitel 4.2 experimentell getestet. Die Untersuchung wird im eng definierten Kontext „Strafen und Anreize im Airline-Bereich“ durchgeführt. Sie beinhaltet die Gegenüberstellung der annäherungsweise identischen Szenarien aus der qualitativen Empirie (vgl. Kap. 3.8) in einer *Online*- und einer *Offline*-Gruppe (vgl. Harris et al., 2006). Die Online-Gruppe entspricht der Kundensteuerung mit *Anreiz-Optionen* und die Offline-Gruppe entspricht der Kundensteuerung mit *Richtlinien*. In beiden Gruppen werden die Probanden mittels der schriftlichen Szenariotechnik (vgl. Kap. 5.2.1) darauf hingewiesen, dass sie den Flug über das Internet buchen. Die unabhängigen Variablen werden im Rahmen des experimentellen Designs mittels der schriftlichen Szenariotechnik manipuliert.

Online-Gruppe

Um die Kundensteuerung und damit die gleichzeitige Untersuchung von Strafen sowie Anreizen durchzuführen, eignet sich der Internet-Buchungsprozess einer Airline (vgl. Kap. 3.8). Anhand der simulierten Online-Buchung kann neben der Hauptvariablen *Kundensteuerung* und deren *Kommunikation* auch der *Ticketpreis* untersucht werden. Für die Pull-Massnahmen ermöglichen auf dem Internet angebotene *Anreiz-Optionen* während des Buchungsprozesses die intendierte Manipulation. Gleichzeitig können mit dem Internet-Buchungsprozess die Push-Massnahmen (Kundensteuerung über *Richtlinien*) initiiert werden, indem die Tarifbedingungen und daraus resultierende Strafgebühren über das Internet kommuniziert werden. Die *Kommunikation* als unabhängige Variable wird somit über die Umsetzung der Kundensteuerung auf der Internetseite der Airline untersucht.

Offline-Gruppe

Die Kundenreaktion auf eine Strafe als Folge der Kundensteuerung über Richtlinien lässt sich am Check-In untersuchen (vgl. Kap. 3.8). Der Check-In ist der erste Kontaktpunkt des Kunden mit der Airline (resp. dem Vertreter der Airline) am Reisetag. Der weitere Verlauf des Airline-Serviceprozesses wird am Check-In definiert (vgl. Kap. 3.1.2). Dies erleichtert die gleichzeitige

Untersuchung verschiedener Push-Massnahmen, die vom Passagier als Strafe wahrgenommen werden können. Indem bei der Befragung verschiedene Strafszenarien präsentiert werden, können die kognitive und emotionale Reaktion sowie das Verhalten der Passagiere analysiert werden.

4.3.2 Methodenwahl

Das in Kapitel 4.2 erarbeitete Forschungsmodell geht von einem Haupteffekt der Kundensteuerung auf die Kundenwahrnehmung und Reaktion aus. Zur Gegenüberstellung der Online- und Offline-Gruppe, und für die Untersuchung der Kausalzusammenhänge unter gleichzeitiger Kontrolle der Störvariablen, eignet sich eine experimentelle Untersuchung (vgl. Kap. 5.1.1; Huber et al., 2006). Die experimentelle Versuchsanordnung hat sich in verschiedenen Untersuchungen in der Forschung zum Dienstleistungsmarketing und Kundenverhalten bewährt (vgl. Estelami und Bergstein, 2006; Mattila und Wirtz, 2006; Levesque und McDougall, 2000; O'Reilly und Puffer, 1989).

Mit der experimentellen Versuchsanordnung können die Einflussfaktoren aus Sicht des Anbieters zielgerichtet manipuliert und die Auswirkungen verschiedener Faktorstufen (vgl. Kap. 5.2.2) auf die aktuellen Prozesse sowie Verhaltensabsichten von Passagieren analysiert werden. Derartige Beziehungen zwischen bestimmten Bedingungen (unabhängige Variable) und den resultierenden Ereignissen (abhängige Variable) lassen sich selten durch reine Beobachtung feststellen (Sarris, 1999; Hicks, 1993).

Abbildung 4-2: Experimente als Grundlage der quantitativen Empirie



Legende: Prov. = provisorisches

Quelle: Eigene Darstellung

Eine experimentelle Untersuchung mit fünf zu manipulierenden Variablen ist umfangreich und stellt hohe Anforderungen an die statistischen Voraussetzungen sowie die Kontrolle der Manipulation (Maxwell und Delaney, 1990). Es ist nicht auszuschliessen, dass bei der Durchführung mit einem Experiment unerwünschte Interaktionen zwischen den unabhängigen Variablen entstehen (Variablen-Konfundierung, vgl. Kap. 5.1.1). Etabliert haben sich in der experimentellen Forschung zum Dienstleistungsmarketing und Kundenverhalten Designs mit maximal vier unabhängigen Variablen (Mattila und Wirtz, 2006; Hamlin, 2005; Wirtz und Chew, 2002). Die quantitative Untersuchung des Forschungsmodells wird daher aufgeteilt in zwei Experimente, die sich bezüglich der Versuchsanordnung unterscheiden (vgl. Abb. 4-2).

Die Versuchsanordnung ist aufgeteilt nach Untersuchungskontext: Zuerst wird ein schriftliches, papierbasiertes Experiment am Flughafen durchgeführt (Experiment I), um die Hypothesen 1 bis 7 zu testen. Anschliessend wird ein web-unterstütztes Experiment (Experiment II) am Heimcomputer der Probanden durchgeführt, um die Hypothesen 8 bis 10 zu testen und Hypothesen 1, 2 und 7 zu validieren. Die Hypothese 10 folgt aus einer Erkenntnis des ersten Experiments (vgl. Kap. 6.4.3). Die Ergänzung der Flughafen-Umfrage mit einem zweiten Experiment erlaubt die:

- Ausweitung der Untersuchung auf verschiedene Populationen und Kulturkreise (Flugpassagiere und Internet-Benützer);
- Kombination der Vorteile verschiedener Erhebungsinstrumente (Kontrolle der Störfaktoren);
- Validierung des Forschungsmodells bezüglich der zentralen Elemente *Kundensteuerung* (h1), *Kommunikation* (h2) und *Verhaltensabsichten* (h7);
- Prüfung der Hypothese 2 (Kommunikation) mit einer manipulierten Website, die einen fiktiven Online-Buchungsprozess visualisiert (vgl. Kap. 5.2.2);
- Prüfung der Hypothese 8 (Einfluss des Geschäftsmodells);
- Prüfung der Hypothese 9 (Einfluss der *Akzeptanz* auf das wahrgenommene *ex post Airline-Image*);
- Prüfung von neuen Erkenntnissen aus dem ersten Experiment (Hypothese 10);
- vertiefte Analyse einzelner Szenarien im Airline-Bereich.

4.3.3 Erkenntnis- und Forschungsobjekt

Das Erkenntnisobjekt der quantitativen Empirie ist das in diesem Kapitel aufgestellte, A-priori-Forschungsmodell mit den dazugehörigen Hypothesen 1 bis 9 und den Nebenhypothesen. Im Mittelpunkt steht die Kundensteuerung mit Push- oder Pull-Massnahmen und deren Auswirkungen auf die Kundenwahrnehmung sowie Verhaltensreaktion. Den theoretischen Bezugsrahmen bilden die disziplinenübergreifenden Betrachtungen aus dem theoretisch-konzeptionellen Teil (vgl. Abb. 4-3).

Flughafen-Experiment

Im ersten, schriftlichen Experiment am Flughafen ist das Forschungsobjekt das Individuum in der Rolle als Flugpassagier. Auf die Grundgesamtheit *Flugpassagiere* fallen Lokal- und Transitpassagiere jeder Flugklasse und Airline am Flughafen Zürich (Unique). Die schriftliche Umfrage vereinfacht den Probanden die Beantwortung des hypothetischen Sachverhalts, da sie direkt am Flughafen zum Airline-Serviceprozess befragt werden (vgl. Kap. 5.2.1).

Ein dem *Feldexperiment* ähnliches Setting weist zwar eine höhere externe Validität auf, ist jedoch anfälliger für Störfaktoren (vgl. Kap. 5.1.1; Winer et al., 1991). Bei einer Befragung am Flughafen, vor dem Boarding, befinden sich die Passagiere in der Airline-Servicekette. Sie haben Zeit und Ruhe für das Ausfüllen des Fragebogens: Die Störfaktoren werden minimiert. Die im Fragebogen vorgestellten Szenarien betreffen zudem Service-Stufen, welche Passagiere vor dem Boarding bereits hinter sich und somit unmittelbar erlebt haben (vgl. Kap. 3.8). Im Vergleich zur Umfrage in der realitätsnahen Umgebung, welche hier angewendet wird, würde ein Laborsetting die Szenarien des Fragebogens abstrakter erscheinen lassen.

Abbildung 4-3: Aufbau und Durchführung der quantitativen Empirie

	Forschungs- und Erkenntnisobjekt	Forschungsziele	Erhebungs- und Auswertungsmethodik
Experiment I (Flughafen)	<ul style="list-style-type: none"> – Flugpassagiere (Lokal- und Transitpassagiere am Flughafen Zürich) – Forschungsmodell: Kundensteuerung als zentrales Element 	<ul style="list-style-type: none"> – Hypothesen h1-h7 testen – Nebenhypothesen testen (Demografie, Interaktion) – Forschungsmodell prüfen 	<ul style="list-style-type: none"> – Experiment am Flughafen, schriftlicher Fragebogen – Deskriptive und konfirmatorische (statistische) Datenanalyse
Experiment II (Online)	<ul style="list-style-type: none"> – Internet-Benützer mit Flugerfahrung – Forschungsmodell: Kundensteuerung als zentrales Element 	<ul style="list-style-type: none"> – Validierung h1, h2 und h7 (h2 mit Web-Manipulation) – Validierung Modell und allfällige Anpassung – Hypothesen h8-h10 testen 	<ul style="list-style-type: none"> – Experiment mit web-gestütztem Fragebogen – Deskriptive und konfirmatorische (statistische) Datenanalyse

Quelle: Eigene Darstellung

Online-Experiment

Mit der Online-Umfrage wird die untersuchte Population erweitert. Im Gegensatz zur schriftlichen Flughafen-Umfrage stehen Merkmalsträger im Vordergrund, die Erfahrung mit Online-Buchungen haben, sowie Passagiere mit Erfahrung auf anderen Flughäfen und aus anderen Kulturkreisen. Durch das Auswahlverfahren mit E-Mail-Versand und durch die Aufschaltung in internationalen Newsgroups werden diese Anforderungen an die Merkmalsträger erfüllt. Die web-gestützte Umfrage findet nicht in der natürlichen Umgebung der Merkmalsträger statt. Störfaktoren wie Lärm oder Versuchsleitereffekte können besser kontrolliert werden (vgl. Kap. 5.1.1; Janetzko et al., 2002).

4.3.4 Forschungsziel

Mit zwei Experimenten sollen das zentrale A-priori-Forschungsmodell und die entsprechenden Hypothesen zur Beantwortung der Forschungsfragen überprüft werden. Zusätzlich wird der Einfluss der demografischen Variablen sowie der Interaktionseffekte im Modell als Nebenhypothesen überprüft. Die postulierten Kausalzusammenhänge werden unter Verwendung entscheidungsstatistischer

Methoden bestätigt oder falsifiziert. Neue Erkenntnisse aus dem ersten (Flughafen-) Experiment sowie die zentralen Kausalzusammenhänge (h1, h2 und h7) sollen mit dem zweiten (Online-) Experiment validiert werden.

Flughafen-Experiment

Im ersten Experiment werden im Sinne der konfirmatorischen Empirie die Hypothesen 1 bis 7 sowie die Nebenhypothesen getestet (Einfluss der Demografie sowie Interaktionen). Zusätzlich zu den Hypothesen wird das Forschungsmodell für die unabhängigen Variablen 1 - 4 anhand einer Analyse der Korrelationsbeziehungen auf neue Zusammenhänge untersucht (vgl. Kap. 6.4.1).

Online-Experiment

Mit dem zweiten Experiment werden allfällige neue Zusammenhänge aus dem ersten Experiment getestet sowie die zentralen Wirkungszusammenhänge im Forschungsmodell repliziert (h1, h2 und h7). Dabei wird durch eine Manipulation des Web-Interfaces bei einem simulierten Buchungsvorgang eine realitätsnähere Untersuchung zur Variablen *Kommunikation* möglich (vgl. Kap. 5.2.2).

Zusätzlich werden im zweiten Experiment die Hypothesen 8 bis 10 erstmalig getestet, wobei Hypothese 10 einen neu untersuchten Zusammenhang im Modell betrifft (vgl. Kap. 6.4.3).

4.4 Zusammenfassung

Das zentrale A-priori-Forschungsmodell mit neun Ursache-Wirkung-Beziehungen wird mit einem Experiment direkt am Flughafen und mit einem web-unterstützten Experiment getestet. Fünf unabhängige Variablen und fünf abhängige Variablen wurden basierend auf den theoretisch-konzeptionellen Grundlagen und der qualitativ-exploratorischen Empirie ausgewählt. Die Zusammenhänge zwischen den Variablen werden im zentralen Forschungsmodell als Hypothesen formuliert. Die endogenen demografischen Variablen sind nicht im Modell abgebildet, da sie nicht experimentell getestet werden. Nach der Flughafen-Umfrage wird das Forschungsmodell für die Online-Umfrage adaptiert und mit einer neuen Wirkbeziehung (Hypothese) geprüft.

Das Modell und die dazugehörigen Hypothesen werden aufgrund der Forschungsfragen aufgeteilt in *Faktoren im Einflussbereich der Airline* und in die *aktuellen Prozesse* sowie *Verhaltensabsichten*. Das Forschungsziel der quantitativen Empirie ist die Bestätigung des Forschungsmodells im Kontext „Strafen und Anreize im Airline-Bereich“ anhand eines Hypothesentests. Hierzu werden im nächsten Kapitel die konzeptionellen Grundlagen zur quantitativen Forschung erarbeitet.

5 Empirisch-konzeptionelle Grundlagen

Für die Analyse des Forschungsmodells und zum Testen der Hypothesen wurde das wissenschaftliche Experiment als geeignete Methode ausgewählt. Die Forschungsmethode beeinflusst die Entwicklung des Fragebogens (Bortz, 2006; Schnell et al., 1999). Sinngemäss werden in diesem Kapitel zuerst die Grundlagen zur experimentellen Untersuchungsmethode für den quantitativen Teil (Kap. 5.1) und dann die Entwicklung des Fragebogens vorgestellt (Kap. 5.2). Anschliessend wird in Kapitel 5.3 die Datenerhebung der vorliegenden Untersuchung beschrieben. In Kapitel 5.4 folgt die Beschreibung der Programme zur statistischen Auswertung.

5.1 Grundzüge der Experimentalforschung

In diesem Kapitel werden die Validitätskriterien des Feld- und Laborexperiments erläutert und auf die vorliegende Untersuchung angewandt (Kap. 5.1.1). Beim web-gestützten Experiment handelt es sich um eine Erweiterung des Laborexperiments. Es wird im Kapitel 5.1.2 vorgestellt. Der Versuchsplan und die Versuchsanordnung des vorliegenden, *strengen* Experiments werden in Kapitel 5.1.3 erläutert. Mit Kapitel 5.1.4 wird aufgezeigt, wie Reaktanz bei den Probanden verhindert wird, wobei die Exploration eine wichtige Rolle einnimmt (Kap. 5.1.5). Dieses Kapitel 5.1 schliesst mit der Berechnung der (*ex ante*) optimalen Stichprobengrösse (Kap. 5.1.6). Im Anhang sind die Grundzüge eines sozialwissenschaftlichen Experiments erläutert (vgl. A.2.2).

5.1.1 Validitätskriterien für Experimente

Die Art des Experiments beeinflusst dessen Validität. Bezüglich der experimentellen Umgebung werden zwei Arten unterschieden: Labor- und Feldexperiment (Westermann, 2000). In einem *Laborexperiment* kann eine Vielzahl an Störvariablen kontrolliert werden, um auf diese Weise die Abhängigkeit der gemessenen Variablen von verschiedenen Bedingungen nachzuweisen. Im *Feldexperiment* hingegen bildet die natürliche Umgebung den Untersuchungsrahmen: Die experimentellen Bedingungen werden in einer natürlichen Umgebung vom Experimentator manipuliert (Hussy und Jain, 2002). Aus methodischer Sicht ist der wesentliche Unterschied zwischen Labor- und Feldexperiment das Ausmass der Kontrolle von Störvariablen (z.B. Lärm, Einflüsse des Versuchsleiters, ungenügende Manipulation der UV). Durch die hohe Kontrolle der Störvariablen werden jedoch im Laborexperiment teilweise Bedingungen geschaffen, die eine Generalisierung auf eine konkrete Situation nicht erlauben (Winer et al., 1991). Feldexperimente sind häufiger situationsrepräsentativ. Sie verfügen über eine höhere externe Validität (im Falle gleich hoher interner Validität) im Vergleich zum Laborexperiment (Sarris und Reiss, 2005).

Cook und Campbell (1979) formulieren vier Arten von Kriterien, die als Massstab zur Beurteilung der Güte eines Experiments verwendet werden können:

Konstrukt- und inferenzstatistische Validität

Ein Experiment verfügt über eine hohe *Konstruktvalidität*, wenn die unabhängige Variable in konzeptuell-theoretisch eindeutiger Weise auf die abhängige Variable wirkt. Dabei liegt keine Konfundierung mit irrelevanten Variablen vor (Kromrey, 2006; Stier, 1999). Die unabhängige Variable als auch die abhängige Variable müssen derart operationalisiert sein, dass sie die im Forschungsmodell eingesetzten Konstrukte tatsächlich messen. Die Konstruktvalidität ist umso höher, je mehr die unabhängigen und abhängigen Variablen das theoretische Konzept tatsächlich repräsentieren (Diekmann, 1998; Hicks, 1993).

Das Kriterium der *inferenzstatistischen Validität* bezieht sich auf die Gültigkeit der Schlussfolgerungen des in einer Stichprobe erhobenen Befundes auf die zugehörige Population (Gesamtheit aller Individuen). Es geht um die Übertragbarkeit eines Effekts der experimentellen Behandlung auf die Population (Girgensohn-Marchand, 1999).

Um die Konstruktvalidität zu gewährleisten, wird in der vorliegenden Untersuchung basierend auf dem Forschungsmodell und Erkenntnissen des theoretisch-konzeptionellen Teils sowie der qualitativen Empirie operationalisiert. Durch Einsatz geeigneter statistischer Methoden (z.B. Entscheidungsstatistik zur Modellprüfung anstatt reine Korrelationsprüfung, vgl. Hirsig, 2001; A.3) wird die inferenzstatistische Validität in der vorliegenden Untersuchung gewährleistet.

Interne und externe Validität

Ein Experiment wird intern valide, wenn die Messwerte der abhängigen Variable eindeutig auf die Manipulation der unabhängigen Variable zurückzuführen sind (Hicks, 1993). Eine notwendige Voraussetzung einer kausalen Interpretation von Effekten des experimentell nachgewiesenen Wirkungszusammenhangs besteht darin, dass keine wissenschaftlich plausible alternative Bedingung für das Zustandekommen des Effekts angeführt werden kann (Campbell und Stanley, 1966). Daten fluktuieren auch ohne Einfluss der experimentellen Bedingung (Birnbaum, 2000). Ursache hierfür sind Störvariablen, die entweder systematisch oder unsystematisch auf die abhängige Variable (Fehlervarianz) einwirken (Hussy und Jain, 2002). Je besser die Störvariablen in einem Experiment kontrolliert werden können, umso höher ist die interne Validität. Um die interne Validität eines Experiments zu erhöhen, formulierte Kerlinger (1973) das *Max-Kon-Min-Prinzip* der Versuchskontrolle:

- Signal (Primärvarianz) maximieren: Wahl von Extremgruppen, Wahl von optimalen Stufen, Umwandlung eines Störfaktors in eine unabhängige Variable;
- Sekundärvarianz (systematischer Fehler) kontrollieren: Eliminierung, Konstanthaltung, Randomisierung, statistische Kontrolle mit Kovarianzanalyse;
- Rauschen (Fehlervarianz) minimieren: Wahl eines Wiederholungs- oder Blockversuchsplans, Untersuchungssituation standardisieren, Zuverlässigkeit des Messinstruments erhöhen.

Ein Experiment, das alle aufgeführten Kriterien berücksichtigt, liesse sich empirisch nicht durchführen (Birnbäum, 2000). In der vorliegenden Untersuchung wird das Signal durch den Einsatz ausgeprägter, dichotomer Faktorstufen maximiert, die Sekundärvarianz durch Randomisierung bei der Probandenauswahl kontrolliert und die Fehlervarianz aufgrund der Exploration und standardisierten Situation minimiert (vgl. Kap. 5.1.3; 5.1.5; 5.2).

Die externe Validität hängt davon ab, inwieweit von der speziellen Versuchssituation auf andere (natürliche) Situationen geschlossen werden kann. Während sich die Konstruktvalidität auf die Variablen bezieht, meint die externe Validität die Gültigkeit der gesamten Versuchsanordnung (Sarris und Reiss, 2005; McGuigan, 1993).

Die Kombination eines Labor- und „Feldexperiments“ (Online- und Flughafen-Experiment) erhöht nicht nur die interne, sondern auch die externe Validität in der vorliegenden empirischen Untersuchung (vgl. Kap. 5.3).

5.1.2 Web-unterstütztes Experiment

Mit Internet-Experimenten können Resultate von Feldexperimenten validiert werden (Frick et al., 2001). Sie sind eine Erweiterung der computerisierten Laborexperimente (Sarris und Reiss, 2005). Durch den kombinierten Einsatz der beiden Erhebungsmethoden wird die Güte der vorliegenden Untersuchung erhöht (vgl. O'Reilly und Puffer, 1989). Bei der Flughafen-Umfrage handelte es sich um eine dem *Feldexperiment* nahe stehende Versuchsanordnung, da die „natürliche“ Umgebung eines Passagiers der Flughafen ist. Das Internet-(Online) Experiment ist kein reines Laborexperiment, aber es erfüllt durch die bessere Kontrolle der Störvariablen eine wichtige Voraussetzung des Laborexperiments (Musch und Reips, 2000). Somit wird ein Teil der Fehler, welcher zu Lasten eines Versuchsleiters geht, vermieden (Reips, 2002). Es wird angenommen, dass sich der Proband bei der Online-Beantwortung der Fragen in einer Umgebung mit wenigen Störvariablen befindet.

Die Versuchsteilnehmer nehmen bei web-gestützten Experimenten mit ihrem Computer über die grafische Oberfläche des Web-Browsers Kontakt zum *Server* auf (vgl. Kap. 5.3.2). Experiment-Generator-Programme wie Superlab (o.V., 2007c) oder WEXTOR (o.V., 2007d) erleichtern die Entwicklung und Veröffentlichung von Experimenten. Mit diesen Programmen erstellte Experimente können programmunabhängig auf beliebigen Computern, die Zugriff aufs Internet haben, durchgeführt werden. Somit lassen sich Experimente an verschiedenen Orten günstig und weitgehend unter denselben Bedingungen durchführen (Frick et al., 2001; Birnbäum, 2000). Die Eingaben der Probanden zeichnet der Web-Server mit einem *Log-File* auf und beantwortet sie (vgl. Reips und Stieger, 2004).

Durch die heute sehr hohe Internetanschlussquote ist eine grosse Anzahl der Allgemeinpopulation für die Experimentalforschung zugänglich (Reips, 2002). Die herkömmliche Stichprobenbeschränkung kann dadurch teilweise umgangen werden. Web-unterstützte Experimentalforschung

kann somit Angehörige bisher vernachlässigter Bildungs- und Altersschichten wie auch Personen aus unterschiedlichen Kulturkreisen mit einbeziehen (Sarris und Reiss, 2005). Zusätzlich zu den oben genannten Vorteilen von webbasierten Experimenten führt Birnbaum (2000) auf:

- Die Teilnahme am Experiment ist nicht ortsgebunden;
- Die Teilnehmerzahl kann potentiell sehr gross und die Erhebungszeit kürzer als im Labor sein, da ohne zeitliche und geografische Beschränkung gleichzeitig viele Personen teilnehmen;
- Es besteht vollkommene Teilnahmefreiwilligkeit. Die Versuchsteilnehmer können jederzeit und ohne Begründung das Experiment abbrechen.

Folgende Nachteile stehen dem gegenüber (Janetzko et al., 2002; Birnbaum, 2000): Die anonyme Versuchsteilnahme führt zu einer geringen Kontrolle und Standardisierung des Verhaltens von Teilnehmern, technische Einflüsse können das Antwortverhalten verzerren oder zu einer Stichprobenbeschränkung führen und die Bereitschaft, an einer bestimmten Untersuchung teilzunehmen, kann mit der untersuchten Variablen konfundiert sein (Selbstselektion). Mit der Durchführung eines Pretests für die Online-Untersuchung sowie der Einbindung des Manipulationschecks werden diese Nachteile im vorliegenden Online-Experiment vermindert und kontrolliert (vgl. Kap. 5.1.5).

5.1.3 Experimentelle Versuchspläne

Für ein *strenges* Experiment müssen drei Gütemerkmale vorausgesetzt werden (Westermann, 2000; Campbell und Stanley, 1966): (1) vor Versuchsbeginn ist eine kausaltheorietische Hypothese vorhanden; (2) die UV ist manipulierbar und geht der AV zeitlich voraus (weitere Bedingungen sind nicht wirksam oder durch Randomisierung ausbalanciert) und (3) die Versuchsbedingungen sowie Störvariablen sind kontrollierbar (bei der Messung besteht höchstens ein Zufalls-, jedoch kein systematischer Fehler).

In der vorliegenden Untersuchung handelt es sich um ein strenges Experiment. Die drei Voraussetzungen sind erfüllt: Das Forschungsmodell und die Hypothesen wurden im Hinblick auf das Experiment entwickelt, wobei auf die Manipulation der UV Rücksicht genommen wurde. Mit einem Versuchsplan im *Between-Subject-Design* (vgl. unten) wird die Randomisierung berücksichtigt und die Störvariablen werden kontrolliert (vgl. Kap. 5.1.1).

Verschiedene experimentelle Versuchspläne werden bei strengen Designs unterschieden (vgl. Sarris, 1999; McGuigan, 1993). An dieser Stelle wird näher auf den Versuchsplan mit Randomisierung (Zufallsgruppenbildung) eingegangen, da dieses Design in der vorliegenden Arbeit verwendet wird.

Versuchsplan mit Randomisierung

Versuchspläne mit Randomisierung werden auch als *Between-Subject-Designs* bezeichnet, da sie auf dem Mittelwertsvergleich zwischen verschiedenen experimentell manipulierten Gruppen basieren (Hamlin, 2005; Hicks, 1993). Eine Gesamtgruppe wird dabei den einzelnen experimentellen Bedingungen (Zellen) randomisiert zugeordnet (Kirk, 1995). Aus der Stichprobentheorie folgt, dass jedes Mitglied einer Gruppe die gleiche Chance haben muss, aus der zuvor definierten Population gezogen zu werden (Stier, 1999). Merkmale der Versuchsteilnehmer, die als Störvariable wirken könnten, verteilen sich somit zufällig auf die einzelnen Bedingungen. Dadurch üben Organismusfaktoren (vgl. Kap. 2.3.1) keinen systematischen Effekt auf die AV aus (Birnbbaum, 2000). Dies erhöht die interne Validität des Experiments und macht eine explizite Kenntnis der Störvariablen unnötig (Hussy und Jain, 2002). Nachteil des *Between-Subject-Designs* sind die Randomisierung, die gemäss dem Gesetz der grossen Zahl eine grosse Stichprobe erfordert (Ökonomie-Problem), und ein relativ hoher Zufallsfehler (vgl. Kap. 5.1.6). Die praktische Umsetzung des vorliegenden Versuchsplans ist in Kapitel 5.3 erläutert.

Versuchsanordnung

Zur Überprüfung von Hypothesen in einem Experiment braucht es mindestens zwei zweistufig (dichotom) angelegte unabhängige Variablen (Hamlin, 2005; Kirk, 1995; Hicks, 1993). Dies bedeutet, es gibt mindestens eine UV A mit der Faktorstufe A1 und A2 sowie eine UV B mit der Faktorstufe B1 und B2. Um alle möglichen Kombinationen empirisch zu testen, sind somit vier Zellen nötig. Man spricht von einem 2x2 Design (Hamlin, 2005). Unter *Faktorstufe* wird die Ausprägung der experimentellen Bedingungen verstanden.

Abbildung 5-1: Versuchsanordnung im 2x2x2x2 Design (Flughafen-Experiment)

		UV2 (t)		UV2 (h)	
		UV3 (t)	UV3 (h)	UV3 (t)	UV3 (h)
UV1 (R)	UV4 (t)				
	UV4 (h)				
UV2 (A)	UV4 (t)				
	UV4 (h)				

Legende: t = tiefe, h = hohe Faktorstufe, A = Anreiz, R = Richtlinie;
 Quelle: Eigene Darstellung

Im vorliegenden Flughafen-Experiment werden, basierend auf dem Forschungsmodell und den dazugehörigen Hypothesen, vier experimentelle Bedingungen (UV1 - UV4) mit dichotomen Faktorstufen getestet (2x2x2x2 Design). Um alle möglichen Kombinationen zu prüfen, werden 16 Zellen benötigt (vgl. Abb. 5-1). Für die praktische Umsetzung bedeutet dies, dass 16 verschiedene Fragebogen entwickelt werden müssen, die sich bezüglich der Manipulation mindestens einer Faktorstufe

unterscheiden. Das Online-Experiment wird mit einem 2x2x4 Design mit 16 Zellen (2^4) durchgeführt: Es werden drei unabhängige Variablen (UV1, UV2 und UV5) manipuliert, wobei UV5 (Geschäftsmodell) vier Faktorstufen aufweist (vgl. Kap. 5.1.6).

5.1.4 Reaktive Messeffekte

Je nach Erfahrung und Persönlichkeitsstruktur kann die Teilnahme bei Probanden durch verschiedene kognitiv-emotionale Zustände begleitet sein, was einen beträchtlichen Einfluss auf den zu untersuchenden Prozess hat (Sarris und Reiss, 2005). Bspw. versuchen Probanden in einem Experiment, eine eigene Deutung des Versuchsziels vorzunehmen. In Experimenten kann der Proband das generelle Ziel des Versuchs kennen oder bezüglich des Versuchsziels als auch hinsichtlich einer speziellen experimentellen Aufgabe im Unklaren gelassen werden (Birnbau, 2000; Winer et al., 1991). Verschiedene Abstufungen der Wissentlichkeit sind somit möglich. In der vorliegenden Untersuchung wird die experimentelle Versuchsanordnung aufgrund von Erkenntnissen des durchgeführten Pretests nicht erwähnt (vgl. Kap. 5.2.5). Es wird zwar unvollständig, aber nicht falsch - im Sinne einer Täuschung - instruiert.

Die Beantwortung der Fragen wird in der vorliegenden Flughafen- und Online-Umfrage dem Probanden selbst überlassen. Dadurch können die Versuchsleitereffekte vermindert werden. Versuchsleiter können durch ihre Merkmale reaktive Messwerte auslösen oder inadäquat auf Verhalten von Probanden reagieren (Westermann, 2000). Durch den Verzicht auf Interviewer wird zudem der Aufwand für die geforderte Stichprobengröße erheblich vermindert (Girgensohn-Marchand, 1999; Laatz, 1993). Stellen sich trotz der Pretests Fragen, so können am Flughafen die verteilenden Studenten weiterhelfen. Bei der Online-Umfrage wird eine E-Mail-Adresse zur Kontaktaufnahme bei Fragen bereitgestellt.

Der Unterschied zwischen der papier- und der internetbasierten Befragung besteht hauptsächlich bei der erhöhten Abbruchrate der Online-Befragung (Manfreda und Vehovar, 2002; Birnbau, 2000). Hiskey und Troop (2002, S. 251) bestätigen: *“There is converging evidence that Web-based survey methods result in qualitatively comparable results to traditional survey, even in longitudinal studies.”* Dadurch, dass der Fragebogen in beiden Varianten selbstständig von den Probanden ausgefüllt wird, sind neben der Ausfallquote keine weiteren gravierenden Verzerrungen zu erwarten (Manfreda und Vehovar, 2002). Um Reaktanz vorzubeugen, ist bei der schriftlichen Instruktion darauf zu achten, dass klar, motivierend und für jedermann eindeutig, verständlich kommuniziert wird (Stier, 1999).

5.1.5 Exploration

Um sich gegen unerkannte Störquellen während des Versuchsablaufes abzusichern und damit die Validität des Experimentes zu erhöhen, empfiehlt sich eine Exploration (McGuigan, 1993; Breden-

kamp, 1980). Ein wichtiges Element der Exploration ist der Manipulationscheck im Anschluss an den Versuch (Boshoff, 1997; Kirk, 1995).

In der vorliegenden Untersuchung wird der Manipulationscheck im Fragebogen integriert (vgl. Kap. 5.2.3). Er dient der Überprüfung, ob bei den Probanden die intendierte Manipulation der UV durch die experimentelle Versuchsanordnung realisiert wurde (Sarris und Reiss, 2005). Bei der Gestaltung des Manipulationschecks ist darauf zu achten, dass keine Suggestivfragen gestellt werden, die den Probanden auf eine bestimmte Antwort drängen (Westermann, 2000; Diekmann, 1998). Bei der Skalenkonfiguration des Manipulationschecks in einem Fragebogen sollte daher eine Ausweichkategorie vorhanden sein. Wird die intendierte Manipulation nicht richtig verstanden, wird der Fragebogen zur Auswertung der betreffenden Variable nicht verwendet. Daher unterscheidet sich die Stichprobengrösse, je nach Erfolg der Manipulation, für jede experimentelle Bedingung.

Ebenfalls zur Exploration gehört ein Pretest (Vorexperiment) mit wenigen Probanden. Ziel ist, zu untersuchen, ob die experimentellen Variablen im Einzelfall auch realisiert werden können (Hussy und Jain, 2002; Janetzko et al., 2002). Die Erfahrung während des Pretests erlaubt die Optimierung des Experiments vor der eigentlichen Durchführung: Anpassungen der Operationalisierung von UV und AV (z.B. Stärke des Ausprägungsgrades) sowie allfällige Verbesserung des Fragebogendesigns (Formulierung, Verständnis, Aufbau). Die Exploration für die vorliegende Untersuchung wird in Kapitel 5.2.3 (Manipulationscheck) und in Kapitel 5.2.5 (Pretests) vorgestellt.

5.1.6 Optimale Stichprobengrösse

Da praktisch nie eine ganze Population von Systemen in eine Erhebung mit einbezogen werden kann, wird auf Stichproben zurückgegriffen (Stier, 1999). Um die Erkenntnisse aus der Stichprobe generalisieren zu können, werden die Hypothesen aus Kapitel 4.2 mit entscheidungsstatistischen Verfahren geprüft (Westermann, 2000; Diekmann, 1998). Sie basieren auf der Prüfung von Hypothesen, die postulieren, dass eine Stichprobe aus einer spezifischen Population stammt (vgl. A.3.2; Laatz, 1993). Die Definition der *Population* und das angewandte Verfahren zur Stichprobenwahl bestimmen die statistische Validität des Experiments entscheidend (Kromrey, 2006; Schnell et al., 1999).

Eines der wirkungsvollsten Verfahren für ein hohes Mass an Stichprobenrepräsentativität ist die Zufallsauswahl (vgl. Kap. 5.1.1; 5.1.3). Weitere Faktoren zur Bestimmung der Anzahl Probanden sind *wissenschaftliche Arbeiten* zum ähnlichen Themenkreis oder *statistische Überlegungen* (Kraemer und Thiemann, 1987).

Population

Im Falle von real existierenden experimentellen Effekten will ein Forscher diese mittels eines Minimums an Erhebungsaufwand auch tatsächlich nachweisen (McGuigan, 1993). Es hängt von verschiedenen Faktoren ab, wie gross die Stichprobe sein soll. Ein wichtiger Faktor ist die Grösse der

Population, denn im Hinblick auf die Population werden die Zufallsstichproben gebildet (Winer et al., 1991; Bredenkamp, 1980). In der vorliegenden Untersuchung ist die Population sehr gross (Flugpassagiere und Internet-Benutzer), weshalb die Berechnung der erforderlichen Stichprobengrösse anhand der Population ungenau wird. Bspw. flogen im Jahr 2006 täglich durchschnittlich 53'000 Personen vom Flughafen Zürich, wovon knapp 1% Transitpassagiere waren (o.V., 2006d).

Wissenschaftliche Arbeiten

Experimentelle Forschung im Bereich Dienstleistungsmarketing und Kundenverhalten postuliert für ein Experiment mindestens 20 Probanden pro Zelle (vgl. Ying und Chung, 2007; Estelami und Bergstein, 2006; Mattila und Wirtz, 2006; Hamlin, 2005). In beiden vorliegenden Experimenten werden 16 Zellen getestet, was eine Gesamtstichprobengrösse (N) von 320 Probanden erfordert (vgl. Kap. 5.1.3). Im nächsten Abschnitt wird aus statistischer Sicht validiert, ob mit einer Gesamtstichprobengrösse von $N = 320$ Personen die statistischen Anforderungen an die vorliegenden Hypothesenprüfungen erfüllt werden.

Statistische Überlegungen

Aus statistischer Sicht wird die Stichprobengrösse durch die Anforderungen zum Hypothesentest bestimmt. Kraemer und Thiemann (1987, S. 9) formulieren hierzu: *“On the assumption that the research hypothesis is true, it is necessary that the investigator determines the sample size (N) needed to give this hypothesis a reasonable chance to be correct.”* Aus statistischer Sicht sind daher die Effektgrösse und das gewünschte Fehler-Risiko pro Hypothese relevant. Aufgrund inhaltlicher Überlegungen können die Effektgrösse als auch das α - und β -Fehler-Risiko festgelegt werden (vgl. A.3). In diesem Fall existiert ein optimaler Stichprobenumfang, mit dem zwischen der Nullhypothese (H_0) und der Alternativhypothese (H_1) eindeutig entschieden werden kann (Bortz, 2006). Folgende vier Richtlinien haben Kraemer und Thiemann (1987) bezüglich der Stichprobengrösse (n) definiert: (1) Je stringenter das Signifikanzniveau, umso grösser ist die benötigte Stichprobengrösse (für $\alpha = 1\%$ ist N grösser als für $\alpha = 5\%$); (2) zweiseitige Hypothesenprüfungen verlangen ein grösseres N als einseitige Tests; (3) je tiefer das Differenzmass (Effektgrösse), umso grösser muss N sein und (4) je grösser die Teststärke ($1-\beta$), umso grösser muss N sein (umgekehrt gilt: je kleiner N , umso tiefer ist die Teststärke).

Grundsätzlich ist das Ziel jeder Untersuchung, die Hypothese mit möglichst kleinem α - und β -Fehler-Risiko zu prüfen und dabei eine ökonomische Anzahl Stichproben zu verwenden. Da diese drei Voraussetzungen oft nicht einzuhalten sind, müssen Kompromisse zwischen Teststärke, Aufwand und α -Fehler eingegangen werden (Atteslander, 2000). Zur *A-priori-Definition* der Stichprobengrösse n pro Faktor (UV) werden für das schriftliche Flughafen-Experiment und für das Online-Experiment folgende Voraussetzungen definiert (vgl. Tab. 5-1):

- Signifikanzniveau von 1% (α -Fehler-Risiko), gemäss Kraemer und Thiemann (1987) als strenger Massstab für den Nullhypothesentest;
- Teststärke ($1 - \beta$) von 80%, gemäss Faustregel für Sozialwissenschaften (Hirsig, 2001);
- Effektgrösse von 0.5 für einen mittleren Effekt, der bei einer sozialwissenschaftlichen Untersuchung zu erwarten ist (Hussy und Jain, 2002);
- Zweiseitige Hypothesentests, da es sich bei den meisten in dieser Arbeit getesteten Hypothesen um ungerichtete Hypothesen handelt;
- Allokationsverhältnis zwischen den Experimentalgruppen n_1 und n_2 von 0.7. Zur statistischen Auswertung werden nur Daten von Probanden verwendet, die den Manipulationscheck bestehen. Daher ist pro Faktorstufe keine ausgeglichene Anzahl an Probanden (Stichproben n_1, n_2) zu erwarten. Das Allokationsverhältnis von 0.7 berücksichtigt diese Tatsache.

Für den *t*-Test ergibt die Poweranalyse (Faul et al., 2007a) mit der Software G*Power und den oben aufgeführten Bedingungen ein *n* von total 198 ($n_1 = 116, n_2 = 82$) pro Hypothese und getestete UV (vgl. Tab. 5-1). Wird das Differenzmass (*d*) von 0.5 zu $d = 0.8$ erhöht (starker Effekt), fällt gemäss G*Power-Analyse die benötigte Stichprobe auf $n = 80$ ($n_1 = 47, n_2 = 33$). Als „strengere“ Anforderung bezüglich der Stichprobengrösse und um eine höhere Teststärke zu erreichen, wird in der vorliegenden Untersuchung von einem Differenzmass von $d = 0.5$ ausgegangen und ein *n* von 198 pro Hypothese angestrebt. Für einen varianzanalytischen *F*-Test (ONEWAY ANOVA) ergibt die Berechnung mit G*Power eine optimale Stichprobengrösse für den Vergleich von fünf Gruppen von $n = 75$ (vgl. Tab. 5-1).

Tabelle 5-1: Optimale Stichprobengrösse für *t*- und *F*-Tests (Poweranalyse)

t-Test	Stichprobengrösse			Test (2-seitig)			Effekt	
	n_1	n_2	total	Kritisches <i>t</i>	<i>df</i>	Teststärke	Sig.	<i>d</i>
Input Parameter (Allokation n_1/n_2)	(1	0.7)				.80	.01	0.50
Output Parameter	116	82	198	2.60	196			
F-Test (ANOVA)	Gruppen			Kritisches <i>F</i>				
Input Parameter	5					.80	.01	0.50
Output Parameter	75			3.60	75			

Legende: ANOVA = Analysis of Variance, *d* = Effektstärke, *df* = Freiheitsgrade, *n* = Stichprobengrösse, Sig. = Signifikanz;

Quelle: Eigene Berechnung mit G*Power

In der Versuchsplanung ist pro Zelle ein balancierter Plan anzustreben: Es sind möglichst gleich viele Beobachtungen pro Zelle zu realisieren, damit unkorrelierte Faktoren vorliegen können (Hussy und Jain, 2002). Zudem vermindern gleiche Zellenhäufigkeiten die Bedeutung der Varianzen-Homogenität, eine Voraussetzung der Varianzanalyse (vgl. A.3.3). Mit der verwendeten Statistik-Software SPSS (vgl. Kap. 5.4) wird bei ungleichen Zellengrößen, wie sie aufgrund der Manipulationschecks zu erwarten sind, eine Korrekturformel nach Brown-Forsythe verwendet (vgl. Roth, 1983). Diese Korrekturformel eliminiert bei der Auswertung Ungleichheiten durch verschiedene Zellengrößen.

5.2 Fragebogenentwicklung

Mit Kapitel 5.2.1 werden Überlegungen zur inhaltlichen Gestaltung des Fragebogens gemacht, wobei die Szenario-Methode im Vordergrund steht. Zur Entwicklung des Fragebogens müssen zuerst die Variablen des zentralen Forschungsmodells operationalisiert werden. Unabhängige Variablen als experimentelle Bedingungen werden anhand von manipulierbaren Faktorstufen operationalisiert (Kap. 5.2.2). Die Auswirkung der verschiedenen Faktorstufen auf die abhängigen Variablen werden im Fragebogen mit Rating-Skalen gemessen. Diese Skalenkonfiguration sowie die Kodierung jeder Variablen und des Manipulationschecks sind in Kapitel 5.2.3 aufgeführt. In Kapitel 5.2.4 wird die Struktur des schriftlichen Fragebogens am Flughafen und des web-gestützten Fragebogens einzeln vorgestellt. Zur Exploration gehört ein Pretest, der ebenfalls für beide Experimente separat durchgeführt wird (Kap. 5.2.5). Die Fragebogen der beiden Experimente sind im Anhang aufgeführt (A.4).

5.2.1 Einführung

Beim Design des Fragebogens für die Untersuchung am Flughafen spielte die Verständlichkeit eine wichtige Rolle (Atteslander, 2000). Es sollten keine Fachausdrücke, keine Fremdwörter und keine Abkürzungen verwendet werden (vgl. Stier, 1999). Alle Fragen waren daher kurz gefasst, eindeutig und bezogen sich auf konkrete, einzelne Sachverhalte. Sie sollten die Befragten nicht überfordern, da am Flughafen Probanden aus allen Bildungsschichten mit einbezogen wurden. Der Fragebogen wurde in einer deutschen und einer englischen Version erstellt, da Probanden aus verschiedenen Sprachregionen angesprochen wurden. Um die Probanden zu motivieren und um die Ausfallquote zu verringern (insbesondere im Online-Fragebogen, vgl. Musch und Reips, 2000), wurden unter den an der Umfrage Teilnehmenden ein Europaflug für zwei Personen sowie Büchergutscheine verlost. Eine richtige Bestrafung von Flugreisenden in einem Feldexperiment kann aus forschungsethischen Gründen nicht durchgeführt werden (vgl. Schoefer und Ennew, 2005). In beiden Experimenten (Flughafen und Online) wurde daher die *Projective Role Playing Method* (Szenario-Methode) für die Fragebogenentwicklung eingesetzt. Die Probanden wurden nicht aufgrund bereits erlebter Situationen, sondern durch die Simulation einer künstlichen Situation befragt (Wirtz und Chew, 2002).

Die Probanden wurden darauf hingewiesen, dass es keine „richtigen“ oder „falschen“ Antworten, sondern nur die persönliche, *projektive* Einschätzung der jeweiligen Situation gibt (vgl. A.4).

Die Technik des *Projective Role Playing* setzt eine hohe Kongruenz zwischen der untersuchten Situation und den Erfahrungen der Probanden voraus (Dabholkar, 1996). Daher muss sichergestellt sein, dass die Probanden den Untersuchungskontext kennen und sich in die hypothetische Situation hineinversetzen können (Hui und Bateson, 1991). Aus diesem Grund drängte sich für das erste Experiment eine Befragung der Probanden direkt am Flughafen auf. Im zweiten Experiment, der Online-Befragung, wurde die Erfahrung des Probanden aufgrund des *Seriousness Check* mit der Abfrage der Flugerfahrung überprüft (vgl. A.4.3).

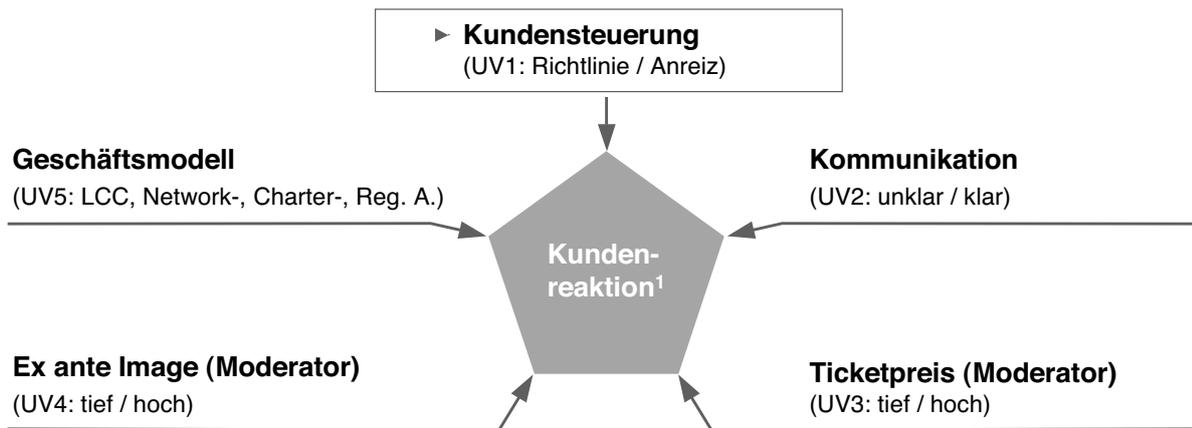
Mit der Szenario-Methode wird die interne Validität erhöht, weil die Probanden nicht nach sozial erwünschten Schemata antworten (Churchill, 1995). Zudem wird das Antwortverhalten standardisiert und nicht durch individuelle Unterschiede in der Verhaltensreaktion oder durch persönliche Umstände verzerrt (vgl. Wirtz und Bateson, 1999; Hui und Bateson, 1991; Havlena und Holbrook, 1986). Schliesslich müssen keine Probanden gesucht werden, die explizit Erfahrung mit einer Bestrafung oder einem Dienstleistungszwischenfall haben, was die Versuchsplanung vereinfacht (Schoefer und Ennew, 2005).

Die Szenario-Methode findet eine breite Anwendung in der Forschung im Bereich des Kundenverhaltens, der Kundenzufriedenheit und der Dienstleistungszwischenfälle (vgl. Wirtz und Mattila, 2001; Levesque und McDougall, 2000; Weiner, 2000; Dabholkar, 1996; Bitner, 1990). Die hohe externe Validität der Methode wurde von Wardman (1988) mit Conjoint-Analysen im Zusammenhang mit Konsumenten-Präferenzen nachgewiesen. Als Nachteil der Szenario-Methode erweist sich die Schwierigkeit der Probanden, sich in die Situation hineinzusetzen. Das Antwortverhalten ist verzerrt, wenn die Probanden nicht so antworten, wie sie sich im richtigen Leben tatsächlich verhalten würden (Wirtz und Chew, 2002). Die Exploration stellt für beide vorliegenden Experimente sicher, dass die Probanden die beschriebene Situation verstanden haben (vgl. Kap. 5.1.5).

5.2.2 Operationalisierung der unabhängigen Variablen

Zur Übersicht ist die auf dem zentralen Forschungsmodell basierende Versuchsanordnung in Abbildung 5-2 dargestellt. Im Experiment am Flughafen werden die unabhängigen Variablen *Kundensteuerung* (UV1), *Kommunikation* (UV2), *Ticketpreis* (UV3) und *ex ante Image* (UV4), im Online-Experiment die Variablen *Kundensteuerung*, *Kommunikation* und *Geschäftsmodell* (UV5) manipuliert.

Als zentrales Element der Untersuchung (neben der Kommunikation) ist die Variable *Kundensteuerung* hervorgehoben. Die Faktorstufen (Manipulationen) sind in Abbildung 5-2 jeweils in Klammern angegeben. Gemessen wird die Kundenreaktion nach jeder Manipulation in den Szenarien A bis D (vgl. Kap. 3.8).

Abbildung 5-2: Experimentelle Versuchsanordnung

¹ **Messen der Kundenreaktion:** Akzeptanz (AV1), negative Verhaltensabsichten (AV2), Fairness-Beurteilung (AV3), Emotion (AV4) und ex post Image (AV5) in vier Szenarien A-D.

Legende: AV = abhängige Variable, UV = unabhängige Variable, Reg. A. = Regional Airline;
Quelle: Eigene Darstellung

Zwei Überlegungen müssen gemäss Westermann (2000) und McGuigan (1993) bei der Wahl der Anzahl Faktorstufen gemacht werden: (1) Zusätzliche Faktorstufen generieren neue Informationen, sie erfordern aber (2) eine erhöhte Anzahl von Versuchspersonen und vermindern gleichzeitig die Teststärke.

Die UVn 1 bis 4 werden in Anlehnung an die Operationalisierung der empirischen Untersuchung zur *Justice Theory* von Schoefer und Ennew (2005) mit dichotomen Faktorstufen manipuliert (vgl. Abb. 5-2). Der Unterschied zwischen den Stufen ist bei der Operationalisierung so zu wählen, dass die Ausprägung der abhängigen Variablen möglichst gross ist (Hussy und Jain, 2002; Sarris, 1999). Mit dichotomen Faktorstufen kann diese Anforderung an die Primärvarianz besser erfüllt werden (Kromrey, 2006). Dichotome Faktorstufen haben sich in der experimentellen Forschung im Dienstleistungsmarketing und in der Forschung zum Kundenverhalten bewährt (vgl. Estelami und Bergstein, 2006; Mattila und Wirtz, 2006; Burns et al., 2000; de Ruyter und Wetzels, 2000).

Da im Airline-Bereich verschiedene Geschäftsmodelle eingesetzt werden (vgl. Kapitel 3.1.1), wird UV5 (Geschäftsmodell) mit vier Faktorstufen im Online-Experiment manipuliert (vgl. Tab. 5-3). Eine dichotome Operationalisierung würde die Ergebnisse der Untersuchung verfälschen.

Im Folgenden wird die Ausgestaltung der Operationalisierung der UVn im Fragebogen für die Flughafen- und Online-Umfrage getrennt aufgeführt, da Unterschiede bestehen.

Flughafen-Umfrage

Die Probanden wurden im schriftlichen Fragebogen mit folgender Ausgangssituation konfrontiert: „Sie planen eine Flugreise, um drei Wochen Ferien an Ihrer Traumdestination zu verbringen. Dieser

Interkontinentalflug in der Economy Class wird zehn Stunden dauern. Sie werden viel Gepäck dabei haben. Da mehrere Airlines die Strecke bedienen, wählen Sie die Fluggesellschaft Y-Air [...]. Sie werden die Reise in 3 Monaten antreten. Ihren Flug buchen Sie über die Internetseite der Y-Air. Zuerst tippen Sie allgemeine Daten wie Startort und Destination sowie das gewünschte Datum ein.“

In dieser Ausgangssituation wurden die unabhängigen Variablen UV2-UV4 manipuliert: Kommunikation (UV2), Ticketpreis (UV3) und *ex ante* Image (UV4).

Das Konzept zur Kundensteuerung (UV1) wurde im Fragebogen durch Anpassung von vier Szenarien an die jeweilige Faktorstufe manipuliert (vgl. Abb. 5-3). Unter der Faktorstufe UV1 „Richtlinie“ (Offline-Gruppe) wurden die Szenarien im Anschluss an die Ausgangssituation als Situationen am Check-In vorgestellt (vgl. Kap. 3.5.5). Die *Anreiz-Optionen* (Online-Gruppe) hingegen werden von einer Airline vorzugsweise während des Online-Buchungsprozesses eingeführt. Daher wurde der Proband für diese Manipulation in allen Szenarien komplett durch die Y-Air Online-Buchung geführt. Zwar „buchte“ der Proband in der Gruppe Faktorstufe „Richtlinie“ online, aber ihm wurden die Szenarien nicht als *Anreiz-Optionen* im Internet, sondern als *Strafen mit Option* am Reisetag vor dem Check-In präsentiert (vgl. Kap. 4.3.1). Daraus ergaben sich folgende Anforderungen an die qualitative Empirie bezüglich der Szenarien-Auswahl für die quantitative Empirie (vgl. Kap. 3.8):

- Das Szenario lässt sich als Anreiz und als Richtlinie gestalten (Manipulation UV1);
- Das Szenario kann als erwerbbarer Anreiz-Option während der Online-Buchung und als Richtlinie mit Strafe am Check-In inhaltlich äquivalent formuliert werden. Beim Check-In (*Straf-Situation*) wird die Tarifbedingung als monetäre Option offeriert. Damit wird die Vergleichbarkeit zwischen den Kundensteuerungsmassnahmen bezüglich der Akzeptanz (AV1) als Kaufverhalten möglich;
- Das Szenario ist praxisnah und für den Probanden gut nachvollziehbar;
- Die Szenarien unterscheiden sich voneinander und betreffen verschiedene Phasen / Bereiche des Airline-Serviceprozesses;
- Mit den Szenarien können das Forschungsmodell und die dazugehörigen Hypothesen getestet werden.

Von den qualitativ evaluierten Szenarien (vgl. Kap. 3.8) haben vier Szenarien alle Anforderungen erfüllt. Abbildung 5-3 gibt eine Übersicht zu den vier Szenarien (A bis D), je nach Faktorstufe der Kundensteuerung (UV1).

Wie aus Abbildung 5-3 ersichtlich, betreffen die Szenarien A bis C monetäre *Anreiz-Optionen* und *Richtlinien*. Die Preise für die verschiedenen Anreiz-Optionen wurden als Durchschnittswerte in der qualitativen Inhaltsanalyse ermittelt und von Experten auf Praxisnähe validiert (vgl. Kap. 3.4; Kap.

Abbildung 5-3: Szenarien für Manipulation der Kundensteuerung (UV1)

	Anreiz-Option „Option“ Online-Buchung	Richtlinie „Strafe“ am Check-In wegen Tarifbedingungen
Szenario A: Übergepäckgebühr	Kauf von 5 kg Übergepäck als Option für CHF 150.-	5 kg Übergepäck: CHF 150.- im Y-Air Office zahlen oder Gepäck am Flughafen lassen.
Szenario B: Lounge-, VIP Boarding	Kauf von Lounge-Zugang und VIP Boarding für CHF 40.-	Kein Zugang Lounge und kein VIP Boarding. Im separaten Y-Air Office zu kaufen (CHF 40.-).
Szenario C: Flexibles Ticket	Flexibles Rückreisedatum als Option: CHF 200.-	Rückreisedatum ändern. Kosten: CHF 200.- (im Y-Air Office zu kaufen).
Szenario D: Business Class Upgrade	Upgrade in Business Class mit Hälfte aller Meilen.	Upgrade aus organisatorischen Gründen nicht möglich (<i>keine Option = strenge Manipulation</i>).

Quelle: Eigene Darstellung

3.6). Da die Faktorstufen „Anreiz“ und „Richtlinie“ bei der Auswertung gegenübergestellt werden, sind die Preise der Optionen unter beiden Faktorstufen gleich. In Szenario D wird nicht monetär geprüft, wobei den Probanden im Fragebogen erklärt wurde, dass sie bereits über eine bestimmte Anzahl an Flug-Meilen verfügen. Zudem ist Szenario D ein „strenges“ Szenario, da es als einziges Szenario im Falle der Richtlinie keine Option beinhaltet.

In der Online-Umfrage wurden lediglich zwei Szenarien geprüft, da die Aufmerksamkeitsspanne von Internet-Probanden tiefer ist (Reips, 2002). Die Auswahl der zwei Szenarien (B und D) basiert auf den Resultaten der statistischen Datenanalyse (vgl. Kap. 6.4.4).

Die experimentelle Bedingung *Kommunikation* (UV2) wurde im Fragebogen anhand der Verständlichkeit und Sichtbarkeit des Hinweises zu den Anreiz-Optionen und Richtlinien während des Buchungsprozesses manipuliert („unklar“ vs. „klar“). Der *Ticketpreis* (UV3) wurde mit zwei verschiedenen Preisen für einen Langstreckenflug in der *Economy Class* manipuliert („tief“ vs. „hoch“). Die Preise für die Variable *Ticketpreis* wurden aufgrund von Durchschnittswerten der qualitativen Analyse verschiedener Airline Websites (vgl. Kap. 3.6) und aufgrund der Ergebnisse des Pretests je nach Fragebogenversion in CHF, EUR oder USD aufgeführt (vgl. Tab. 5-2).

Da *Image* ein multidimensionales Konstrukt ist (vgl. Kap. 2.1.1), wurden zur Manipulation des *ex ante Image* (UV4) verschiedene Dimensionen der Dienstleistungsqualität der Airline beschrieben (vgl. Tab. 5-2): Service an Bord und am Boden, technischer Stand der Flugzeugflotte und das Ergebnis einer Kundenumfrage zur Qualität der Airline. Das Airline-Image beim Kunden wird in Anlehnung an die experimentelle Untersuchung von Huber et al. (2006) mit zwei Faktorstufen ope-

rationalisiert. Die Faktorstufe *ex ante Image* „hoch“ entspricht den *high-quality brands* und die Faktorstufe „tief“ den *low-quality brands* (vgl. Kahneman und Tversky, 1979).

Tabelle 5-2: Manipulation von Kommunikation, Preis und Image (UV2-UV4)

	Kommunikation (UV2)	Ticketpreis (UV3)	Ex ante Image (UV4)
„hoch“	Bevor Sie die Buchung über das Internet abschliessen, klicken Sie auf einen gut sichtbaren Hinweis. Sie lesen die dort deutlich beschriebenen Tarifoptionen und Bedingungen Ihres Tickets.	Economy Class für CHF 3'000.- inkl. Taxen / Gebühren	Sie wissen, dass bei Y-Air der Service an Bord und am Boden sehr gut ist. Die Flugzeugflotte entspricht dem neuesten Stand der Technik. Bei Kundenumfragen schneidet Y-Air regelmässig überdurchschnittlich ab.
„tief“	Bevor Sie die Buchung über das Internet abschliessen, klicken Sie eher zufällig auf einen schlecht sichtbaren Hinweis. Sie überfliegen die dort undeutlich beschriebenen Tarifoptionen und Bedingungen Ihres Tickets.	Economy Class für CHF 400.- inkl. Taxen / Gebühren.	Sie wissen, dass die Y-Air nur durchschnittlichen Service an Bord und am Boden anbietet. Die Flugzeugflotte entspricht nicht mehr dem neuesten Stand der Technik. Bei Kundenumfragen schneidet Y-Air regelmässig unterdurchschnittlich ab.

Quelle: Eigene Darstellung

Online-Umfrage

Wie bereits in Kapitel 4.2.2 erläutert, wurden die experimentellen Bedingungen Kundensteuerung (UV1) und deren Kommunikation (UV2) mit der Online-Umfrage repliziert, da es sich um zentrale Elemente des Forschungsmodells handelt. Um die Teststärke des Online-Experiments nicht zu beeinträchtigen, wurden mit der Kundensteuerung, Kommunikation und dem Geschäftsmodell (UV5) nicht vier, sondern drei unabhängige Variablen manipuliert, da das Geschäftsmodell nicht dichotom skaliert ist (2x2x4 Design, vgl. Hamlin, 2005; Mattila und Wirtz, 2004). Die Variable Geschäftsmodell wurde anhand von vier Geschäftsmodellen untersucht (vgl. Tab. 5-3): *Low Cost Carrier*, *Network Carrier*, *Charter Airline* und *Regional Airline*.

Für die Kundensteuerung und die Kommunikation wurden dieselben dichotomen Faktorstufen wie in der schriftlichen Flughafen-Befragung verwendet. Die Online-Untersuchung ermöglichte eine realitätsnahe Manipulation der Variablen Kommunikation, indem den Probanden je nach Faktorstufe eine andere Y-Air Buchungs-Website präsentiert wurde. In optischer und inhaltlicher Anlehnung an verschiedene Internet-Buchungsseiten von Airlines wurde eine neutrale Website für beide Faktorstufen *Kommunikation* von Y-Air programmiert.

Die Probanden sahen unter der Faktorstufe Kommunikation „hoch“ eine deutliche Darstellung der Anreiz-Optionen oder Tarifbedingungen. Optisch fällt die klare Strukturierung sowie der (farblich) angepasste Text auf der rechten Seite der Website auf (vgl. Abb. 5-4).

Da der Ticketpreis in der Online-Untersuchung nicht im Untersuchungsfokus stand, wurde der Preis des Flugtickets an das manipulierte Geschäftsmodell angepasst (vgl. Abb. 5-4: Auf dem Internet dargestellte Anreiz-Option mit deutlicher (klarer) Kommunikation unter dem Geschäftsmodell *LCC*).

Abbildung 5-4: Web-Manipulation mit deutlicher Kommunikation der Anreiz-Optionen

Buchen Sie Ihren Flug mit Y-Air	
Ticket	Tarifoptionen
Economy Class Y-Air EASY 1 Erwachsener	VIP Boarding: Nein (<i>hier kaufen</i>) Loungezugang: Nein (<i>hier kaufen</i>) Flexible Umbuchung: Nein (<i>hier kaufen</i>) Mehr als 20kg Gepäck: Nein (<i>hier kaufen</i>) Meilenupgrade möglich: Ja (<i>hier einlösen</i>) 
Hinflug: 30. Juli, 2007 08:00 Heimatstadt - 11:00 Destination	
Rückflug: 5. August, 2007 10:00 Destination - 13:00 Heimatstadt	
Gesamtpreis: EUR 100 (Taxen inkl.)	

Quelle: Eigene Darstellung

Im Gegensatz zur Kommunikation „klar“, wurde bei der Faktorstufe „unklar“ mit einer kleineren Schrift undeutlich und unstrukturiert kommuniziert. Ebenfalls angepasst waren die Preise der Tickets (in EUR oder USD), wobei ein realer Durchschnittspreis verschiedener Airlines (vgl. A.5.2) des jeweiligen Geschäftsmodells als Vorlage diente.

Mit Abbildung 5-5 sind auf dem Internet dargestellte Tarifbedingungen mit undeutlicher (unklarer) Kommunikation unter dem Geschäftsmodell *Charter Airline* dargestellt.

Da in der Online-Untersuchung die Variable *ex post Image* im Vordergrund stand, war die gleichzeitige Manipulation der UV *ex ante Image* nicht möglich. Aus modelltheoretischen Gründen kann nicht dieselbe Variable sowohl als abhängige als auch unabhängige Variable in einer Versuchsanordnung experimentell untersucht werden (Westermann, 2000).

Mit der Untersuchung der AV *ex post Image* wird beabsichtigt, die Auswirkung der Kundensteuerungsmassnahme auf das Image zu untersuchen. Da die Manipulation des *ex ante Image* gleichzeitig das gemessene *ex post Image* beeinflusst, würde bei gleichzeitiger Berücksichtigung beider Variablen eine Variablen-Konfundierung auftreten. Der Einfluss der Kundensteuerungsmassnahme auf die Marken-Wahrnehmung wäre nicht eindeutig nachweisbar.

Mit dem nächsten Kapitel werden die abhängigen Variablen und deren Kodierung erläutert (Skalenkonfiguration).

Abbildung 5-5: Web-Manipulation bei undeutlicher Kommunikation der Richtlinien

Buchen Sie Ihren Flug mit Y-Air

Ticket	Übersicht
Economy Class Y-Air EASY 1 Erwachsener	Die angegebenen Preise gelten für den Hin- und Rückflug bei Y-Air. Treibstoffzuschlag und Flughafentaxen sind bereits enthalten. <div style="text-align: right; margin-top: 20px;">  </div>
Hinflug: 30. Juli, 2007 08:00 Heimatstadt - 11:00 Destination	
Rückflug: 5. August, 2007 10:00 Destination - 13:00 Heimatstadt	
Gesamtpreis: EUR 250 (Taxen inkl.)	
Tarifoptionen: <small>Kein separater Check-In. Kein Loungezugang. Flugabsage und Rückerstattung nicht möglich. Kein Umbuchen. Eine Buchung in dieser Tarifklasse kann nicht für ein Upgrade mit Y-Air Meilen verwendet werden. Prozentsatz der Meilengutschrift für Flüge auf Ihrer Reise: 100%. Maximal zulässiges Gewicht und Abmessungen pro Handgepäck: 9 kg / 55 x 40 x 20cm. Maximal zulässiges Gewicht für Check-In Gepäck: 20kg. Nur ein Handgepäck zulässig.</small>	

Quelle: Eigene Darstellung

5.2.3 Skalenzonfiguration

Die abhängigen Variablen (AV) wurden mit der in der Sozialforschung bekannten Skalierungsmethode nach Likert gemessen (Stier, 1999). Die Probanden konnten mit Hilfe sogenannter *Ratings* einer Aussage zustimmen oder nicht. Die Antwortkategorien waren vorgegeben.

Die Skalierung der Daten beeinflusst die verwendbaren statistischen Verfahren (Schnell et al., 1999). In der vorliegenden empirischen Untersuchung wurde für die AV gemäss Hirsig (2001) und Stier (1999) vom Intervallskalenniveau ausgegangen: Die Differenz zwischen zwei aufeinander folgenden Ausprägungsgraden ist gleich gross und der Nullpunkt wird willkürlich festgelegt. Eine verbale Verankerung drängte sich auf, da u.a. emotionale und kognitive Zustände abgefragt wurden (vgl. Machleit und Eroglu, 2000). In allen Szenarien und für beide Experimente wurde dieselbe Skalenzonfiguration für die AV verwendet. Damit konnte das Antwortverhalten über die Szenarien A bis D verglichen und konnten die Forschungsergebnisse des ersten und zweiten Experiments bezüglich des gemeinsamen Forschungsmodells validiert werden (vgl. Malhotra, 2004).

Um die Anzahl Kategorien für die Rating-Skala zu definieren, wird von der Diskriminierungsfähigkeit der Probanden ausgegangen (Jäpel, 1985). Auch die Differenziertheit des befragten Sachverhaltes spielt eine Rolle und die Art der Datenerhebung (schriftlich, mündlich). Stier (1999) und Jäpel (1985) empfehlen generell 5 +/- 2 Skaleneinheiten. In der vorliegenden Untersuchung wurde mit 6er und 7er Likert-Skalen gemessen. Die 6er Skalen wurden für Verhaltensabsichten verwendet und die 7er Skalen für kognitive und emotionale Zustände. Zur optischen Unterscheidung zwischen

Tabelle 5-3: Manipulation des Geschäftsmodells (UV5)

Low Cost Carrier	Network Carrier	Charter Airline	Regional Airline
Als Billigairline bietet Y-Air günstige Preise an, wobei traditionelle Passagierservices nicht oder nur gegen Bezahlung erhältlich sind (z.B. keine Verpflegung an Bord).	Als Netzwerk Airline fliegt Y-Air in einem Hub-System, um gute Flugverbindungen zu ermöglichen. Sie bietet vollen Passagierservice auf internationalen Flügen zu mehr als 120 Destinationen an (z.B. Verpflegung).	Als Charter Airline fliegt Y-Air ausschliesslich zu Feriendestinationen. Zusätzlich offeriert Y-Air „Ferienpakete“ (Flug & Übernachtung).	Als Regional Airline fliegt Y-Air mit kleineren Flugzeugen von ländlichen Gebieten zu internationalen Flughäfen.

Quelle: Eigene Darstellung

den 6er und 7er Skalen war die 7er Likert-Skala jeweils vertikal angeordnet und verbal verankert (vgl. Tab. 5-4).

Die demografischen Variablen zur Person und zur Flugenerfahrung sind in beiden Fragebogen nominalskaliert. Sie basieren auf Grundlagen zur experimentellen Forschung (Reips, 2000) und auf Forschungsarbeiten im Airline-Bereich (Taneja, 2002; Sultan und Simpson, 2000; Bejou und Palmer, 1998). Bei Nominalskalen kann keine Rangreihenfolge aus der Zuordnung gefolgert werden (Diekmann, 1998).

Variablen zur Verhaltensabsicht

Die Antwortkategorien der Variablen *Akzeptanz* (AV1) und *negative Verhaltensabsichten* (AV2) waren auf einer 6er Likert-Skala intervallskaliert. Die gerade Anzahl von Kategorien verhinderte, dass Probanden aus Bequemlichkeit keine Entscheidung treffen. Um der Gefahr einer Verzerrung des Datenmaterials zu entgehen, wurde eine Ausweichkategorie eingeführt („Ich möchte nicht antworten“). Die Intervallskalen waren numerisch verankert, wobei die Endpunkte verbal erläutert waren.

Die *negativen Verhaltensabsichten* wurden basierend auf Erkenntnissen von White und Yu (2005), Andreassen (1999), Söderlund (1998) und Zeithaml et al. (1996) gemessen mit: *Beschwerde*, *Negative Word-of-Mouth* (*neg. WOM*) und *Illoyalität*. Die numerische Skalierung von *Beschwerde* und *neg. WOM* reichte von „Nein, sicher nicht (1)“ bis „Ja, mit Sicherheit (6)“, wobei „Ja“ bedeutet, dass sich der Passagier beschweren würde, resp. Freunden vom Ereignis erzählt. Bei *Illoyalität* hingegen war die Skalierung von „Meine Lust, wieder mit Y-Air zu reisen, ist gesunken (1)“ bis „Lust ... gestiegen (6)“ verankert. Diese verbale Verankerung der Endpunkte wurde gewählt, um den Probanden bei jeder Frage von „minus“ zu „plus“ zu führen. Daher wird bei den statistischen Testhypothesen für *Beschwerde* und *neg. WOM* eine gegensinnige, für *Illoyalität* eine gleichsinnige Korrelation gefordert: Je höher die Akzeptanz, umso höher das Rating bei *Illoyalität* (vgl. Kap. 6.3).

Die *Akzeptanz* wurde mit „Bezahlen Sie die Gebühr?“ als konkretes Verhalten abgefragt (vgl. Cronin et al., 2000). Die 6er Likert-Skala war mit „Nein, sicher nicht (1)“ und „Ja, mit Sicherheit (6)“ an den Endpunkten verbal verankert.

Tabelle 5-4: Übersicht zur Skalenkonfiguration

Variable	Modell- bezeichnung	Skalierung	Verbale Deskriptoren	Numerische Deskriptoren	Grundlage
Akzeptanz	AV1	6er RI A	Nein	Ja ¹	Cronin et al. (2000)
Neg. Verhaltensab.					
Beschwerde	AV2_1	6er RI A	Nein	Ja ¹	White und Yu (2005);
Negative WOM	AV2_2	6er RI A	Nein	Ja ¹	Söderlund (1998);
Illoyalität	AV2_3	6er RI A	Nein	Ja ²	Zeithaml et al. (1996).
Fairness-Beurteilung	AV3	7er RI	Ja	Nein	McColl-Kennedy; Sparks (2003) u.a.
Emotionale Reaktion	AV4	7er RI	Ja	Nein	Oliver und Swan (1989)
Ex post Image	AV5	3er RI	Ja	Nein	Huber et al. (2006) u.a.
Manipulationscheck	-	3er RI	Nein	Ja ³	Sarris (1999)
Demografie	-	N A	Ja	Nein	Reips (2000) u.a.

Legende RI = Rating-Skala / Intervallskalierte Antwortkategorien, N = Nominalskalierte Antwortkategorien, A = Ausweichkategorie „Ich möchte nicht antworten“, Neg. Verhaltensab. = Negative Verhaltensabsichten.

¹ = Endpunkte numerisch verankert: „Nein, sicher nicht“ ... „Ja, mit Sicherheit“

² = Endpunkte numerisch verankert: „Lust...gesunken“ ... „Lust...gestiegen“

³ = Faktorstufe UV „tief“ ... „hoch“ (Manipulationscheck angepasst an UV)

Quelle: Eigene Darstellung

Variablen zur kognitiven und emotionalen Reaktion

Da zur Beurteilung der *Fairness* (AV3) und *Emotion* (AV4) mit neutralen oder indifferenten Antworten zu rechnen war, wurde eine ungerade 7er Likert-Skala eingeführt. Die ungerade Anzahl von Kategorien ermöglichte eine balancierte Skala mit einer neutralen Position. Balancierte Skalen sind nicht in eine Richtung „verzerrt“. Die Anzahl der „positiven“ und „negativen“ Kategorien sind gleich (Girgensohn-Marchand, 1999; Laatz, 1993). Es existiert keine einheitliche Meinung, ob eine gerade oder ungerade Anzahl an Antwortkategorien verwendet werden soll (Stier, 1999). Für jede Auswahlkategorie der 7er Skalen wurden verbale Deskriptoren eingesetzt (vgl. Tab. 5-5).

Die Fairness-Beurteilung war basierend auf Erkenntnissen von McColl-Kennedy und Sparks (2003), Lapidus und Pinkerton (1995), Cropanzano und Folger (1991) sowie der qualitativen Empirie (vgl. Kap. 3.3) verankert mit „sehr unfair (1)“ bis „sehr fair (7)“ und einer neutralen Position. Ausgehend von bewährten Emotionsmessskalen (vgl. Andreassen, 2000; Price et al., 1995; Edell und Burke, 1987) und in Anlehnung an die Likert-Skala von Oliver und Swan (1989) wurde die emotionale Reaktion von „wütend (1)“ bis „glücklich (7)“ verbal verankert.

Manipulationscheck

Im schriftlichen Fragebogen wurde die Manipulation der *Kundensteuerung* (UV1), aufgrund der zentralen Stellung im Forschungsmodell, pro Szenario auf zwei 6er Likert-Skalen getestet. Für jede

Tabelle 5-5: Kodierung der Skalen

Kodierung	1	2	3	4	5	6	7
Verhalten (AV1, AV2)	Nein, sicher nicht	2	3	4	5	Ja, mit Sicherheit	-
Fairness- Beurteilung (AV3)	sehr unfair	unfair	eher unfair	neutral	eher fair	fair	sehr fair
Emotionale Reaktion (AV4)	wütend	unzufrie- den	negativ überrascht	neutral	positiv überrascht	zufrieden	glücklich
Image (AV5)	mässig	neutral	gut	-	-	-	-

Quelle: Eigene Darstellung

Faktorstufe (Richtlinie und Anreiz) wurde der Manipulationscheck durchgeführt. Die Formulierung der Fragen wurde angepasst, um das Antwortverhalten der Probanden nicht zu beeinflussen. Die Probanden wurden gefragt, ob Sie den beschriebenen Sachverhalt als „Zusatznutzen“ oder als „negative Konsequenz“ empfinden. Durch Vermeidung der Begriffe „Strafe“ oder „Anreiz“ konnte eine Beeinflussung beim Antwortverhalten vermieden werden. Im Online-Fragebogen wurde, zur Vereinfachung des Fragebogens, anstatt zwei 6er Likert-Skalen eine 3er Likert-Skala mit einer neutralen, mittleren Position verwendet.

Im Gegensatz zur Kundensteuerung war der Manipulationscheck für die *Kommunikation* (UV2) am Ende des schriftlichen Fragebogens platziert. Nachdem die Probanden alle Fragen zu den Szenarien beantwortet hatten, wurde mit zwei 3-stufigen Likert-Skalen gefragt, ob die Anreiz-Optionen „klar / unklar“ oder „fair / unfair“ kommuniziert waren. Der Manipulationscheck für *Ticketpreis* (UV3) und *ex ante Image* (UV4) wurde mit je einer 3er Likert-Skala im schriftlichen Fragebogen getestet (Preis „hoch / tief“ und Image „hoch / tief“). Für alle aufgeführten Manipulationschecks mit 3er Likert-Skalen wurde als Ausweichkategorie eine mittlere, neutrale Position verwendet. Damit wurde die Teststärke des Manipulationschecks erhöht (Sarris, 1999).

Im Online-Fragebogen wurde zusätzlich die Manipulation der experimentellen Bedingung des *Geschäftsmodells* (UV5) mit einem *Drop-down-Menu* (verbale Deskriptoren) überprüft. Die mit der Website optisch manipulierte Kommunikation wurde mit einer 3er Likert-Skala überprüft („klar / unklar“). Zusätzlich zu den Manipulationschecks wurde in diesem Abschnitt des Online-Fragebogens das Rating des *ex post Image* (AV5) der Airline abgefragt. Hierzu wurde eine 3er, verbal verankerte Likert-Skala verwendet (Image „hoch / tief“). Die Skalierung basierte auf der experimentellen Untersuchung von Huber et al. (2006) sowie LeBlanc und Nguyen (1996). In der Markenforschung werden meist mehrdimensionale Verfahren zur quantitativen Imagemessung eingesetzt (Drenger, 2006). Da die Variable *Image* in der vorliegenden Untersuchung nicht im Vordergrund steht, wird das *ex post Image* mit der oben beschriebenen, eindimensionalen Likert-Skala mit einer Ausweichkategorie gemessen.

5.2.4 Fragebogenstruktur

Flughafen-Umfrage

Der schriftliche Fragebogen für das Experiment am Flughafen war in sieben Teile gegliedert:

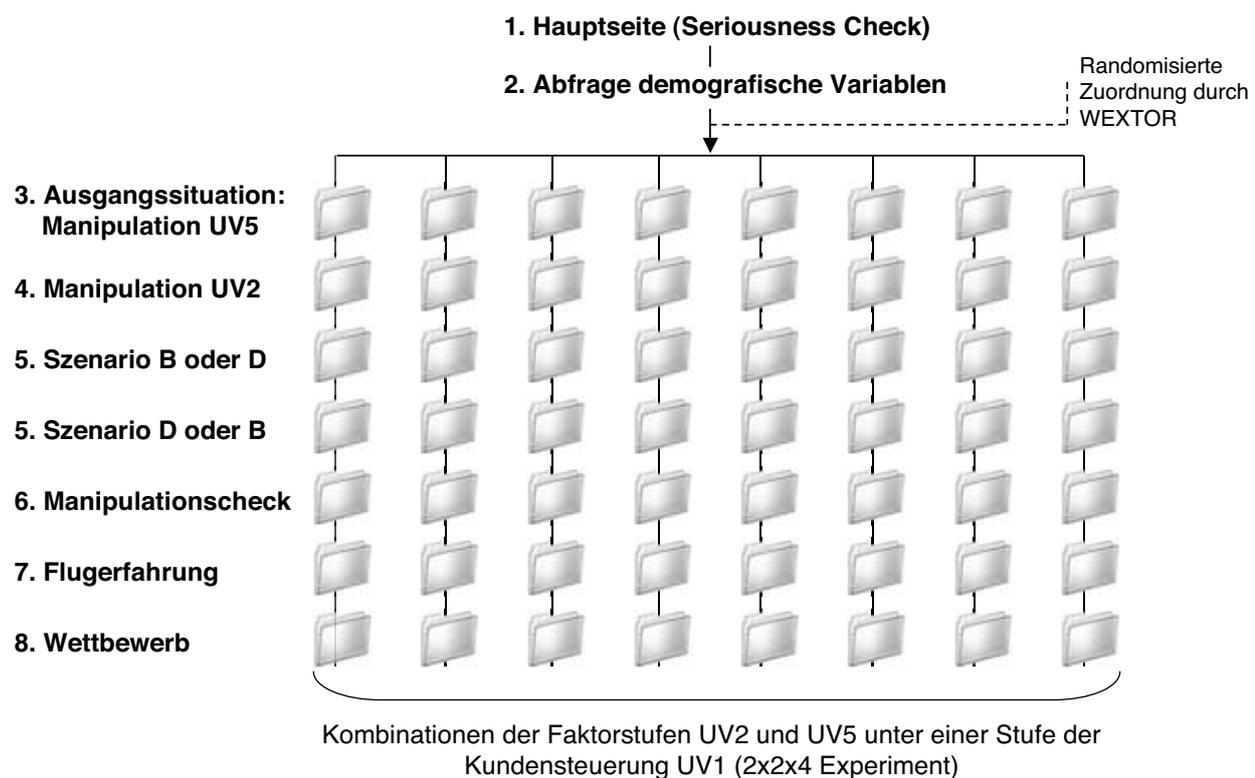
1. In der *Einleitung* wurden Dauer, Zweck und Absender des Fragebogens aufgeführt. Um die Motivation der Probanden zu erhöhen, wurden Flugtickets und Büchergutscheine unter allen Teilnehmenden, die den Fragebogen vollständig ausfüllten, verlost. Den Probanden wurde zudem erklärt, dass alle gewonnenen Daten vertraulich behandelt und anonym ausgewertet werden.
2. Im zweiten Teil des Fragebogens wurden die *demografischen Daten zur Person* abgefragt. Die demografischen Fragen waren am Anfang des Fragebogens platziert, da sich daraus eine höhere Motivation zur Beantwortung der restlichen Fragen ergibt (Frick et al., 2001).
3. Im dritten Teil, der *Ausgangssituation*, wurden die Probanden gebeten, sich in die beschriebene Situation einer Flugreise mit Y-Air hineinzuversetzen: UV2-UV4 wurden manipuliert (vgl. Kap. 5.2.2).
4. Im vierten Teil wurden die *Ausprägungen der abhängigen Variablen* mit den Szenarien A-D abgefragt. Um Verzerrungs- und Platzierungseffekte (Schnell et al., 1999; Stier, 1999) zu vermeiden, war jedes Szenario getrennt auf einer Seite aufgeführt. Die Reihenfolge der präsentierten Szenarien und einzelner Fragen wurden variiert. Damit liess sich verhindern, dass Probanden einzelne Fragen oder Szenarien systematisch zueinander in Beziehung setzen (Girgensohn-Marchand, 1999).
5. Im fünften Teil, dem *Manipulationscheck*, wurden die Probanden gefragt, wie sie die beschriebene Ausgangssituation bewerten. Dieser Abschnitt wurde nicht als Manipulationscheck beschrieben, sondern als „Fragen zur Ausgangssituation“.
6. Da am Flughafen eine heterogene Gruppe von Flugreisenden befragt wurde, wurden die Probanden im sechsten Teil zu ihrem *aktuellen Flug* und zur *Flugerfahrung* befragt.
7. Die letzte, offene Frage erlaubte weitere Einsichten zum Fragebogen: „Welche Gedanken hatten Sie beim Ausfüllen dieses Fragebogens?“ Damit wurden Befragte nicht in eine bestimmte Richtung gelenkt und es konnten zusätzliche, *qualitative Informationen* zum Aufbau und Inhalt des Fragebogens gewonnen werden (vgl. A.5.3). Der Fragebogen endete mit dem Hinweis zur Verlosung eines Fluges und der Abfrage nach Kontaktmöglichkeiten im Falle eines Gewinns.

Online-Umfrage

Die Entwicklung des Online-Fragebogens basierte auf den qualitativen und quantitativen Erkenntnissen des ersten Experiments (schriftliche Umfrage am Flughafen) und auf den Anforderungen an eine web-gestützte Umfrage (vgl. Kap. 5.1.2).

Der Online-Fragebogen folgte demselben Aufbau wie der schriftliche Fragebogen. Er wurde ebenfalls auf Deutsch und Englisch publiziert. Die Online-Befragung wurde bezüglich der Struktur und der experimentellen Versuchsanordnung mit WEXTOR 2.5 programmiert (o.V., 2007d; Reips und Neuhaus, 2002). Anschliessend wurden die einzelnen Seiten mit dem HTML-Editor *Dreamweaver 8* grafisch angepasst (o.V., 2005b). WEXTOR als web-basierte Software wurde gewählt, weil es die Versuchsanordnung des *Between-Subject-Designs* umsetzt und die Validitätskriterien an ein Internet Experiment erfüllt (vgl. Sarris und Reiss, 2005; Birnbaum, 2000). Das *Hosting* des Fragebogens wurde auf einem Rechner des Lehrstuhls für Sozial- und Wirtschaftspsychologie der Universität Zürich durchgeführt. Um auf den Fragebogen zuzugreifen, wurde auf die URL der Hauptseite mit dem Linkgenerator <http://www.tinyurl.com> anonymisiert hingewiesen (vgl. Abb. 5-6).

Abbildung 5-6: Struktur Online-Umfrage



Quelle: Eigene Darstellung

Damit Probanden die Versuchsanordnung nicht durchschauen konnten, war die URL jeder einzelnen Web-Seite des Experiments anonymisiert beschriftet (Reips, 2002). Um die Motivation der Probanden vor Beginn des Online-Experiments zu testen und um eine höhere Datenqualität zu erreichen,

wurde eine *High Hurdle* eingeführt (vgl. Frick et al., 2001). Beim Start der ersten Seite (index.html) wurde im Hintergrund eine virtuelle Datei geladen, welche die Zugriffszeit erhöhte. Damit konnten Probanden ausgesondert werden, die wenig Zeit und Motivation für die Beantwortung des Fragebogens aufwenden wollten (Frick et al., 2001; Birnbaum, 2000).

Die Probanden navigierten von einem Abschnitt (Internetseite) zum nächsten mit einem „Weiter“-*Button*. Mit Hilfe von *Drop-down-Menus* sowie verbalen Deskriptoren und *Buttons* konnten die Probanden die zutreffenden Antworten per Mausklick auswählen. Der Online-Fragebogen basierte auf acht Abschnitten, die jeweils einer Internetseite entsprachen (vgl. Abb. 5-6).

1. In der *Einleitung* wurden die Probanden gefragt, ob Sie an der Studie ernsthaft teilnehmen möchten (*Seriousness Check*) und ob sie bereits einmal geflogen sind. Damit konnte verhindert werden, dass Daten von Personen ausgewertet werden, die den Fragebogen zu Testzwecken ausfüllen oder keinen Bezug zur Thematik haben. Die Probanden wurden in der Einleitung neben Sinn und Zweck der Untersuchung darauf aufmerksam gemacht, dass sie einzelne Fragen bei Bedarf überspringen können.
2. Im zweiten Teil des Online-Fragebogens wurden dieselben *demografischen Daten zur Person* wie im schriftlichen Fragebogen abgefragt. Nachdem die Probanden in diesem Abschnitt „Weiter“ wählten, ordnete WEXTOR den Probanden per Zufallsauswahl den verschiedenen experimentellen Bedingungen zu. In Abbildung 5-6 sind alle möglichen Kombinationen der experimentellen Bedingungen von Kommunikation (UV2) und Geschäftsmodell (UV5) dargestellt. Im 2x2x4 Online Setting sind insgesamt 16 verschiedene Kombinationen (Zellen) möglich (vgl. Kap. 5.1.3). Da die Abbildung 5-6 aus Platzgründen nur unter einer Faktorstufe der Kundensteuerung dargestellt ist (z.B. Anreiz-Option), sind acht verschiedene Durchgänge abgebildet. Jeder Durchgang entspricht der Testanordnung eines Probanden.
3. Im dritten Teil, der *Ausgangssituation*, wurde im Gegensatz zum schriftlichen Fragebogen nicht eine zweiwöchige Ferienreise (Langstrecke), sondern ein mehrtägiger Punkt-zu-Punkt-Flug zu einer „Lieblingsdestination“ (Kurzstrecke) vorgestellt. Damit konnte verhindert werden, dass der Reisezweck einen systematischen Einfluss auf das Antwortverhalten ausübt. Für die Manipulation des Geschäftsmodells (UV5) wurden den Probanden je nach Faktorstufe die verschiedenen Geschäftsmodelle der Y-Air kurz vorgestellt.
4. Um die *Kommunikation* (UV2) zu manipulieren, folgte am Ende der Ausgangssituation ein Hinweis: „Bevor Sie die Buchung über das Internet abschliessen, erscheint folgende Y-Air Webseite“. Damit wurde im vierten Teil die optische Manipulation der Variablen Kommunikation durchgeführt (vgl. Abb. 5-4; 5-5).
5. Im fünften Teil wurden die Szenarien B (Lounge-Zugang) sowie D (Meilen-Upgrade) wie im schriftlichen Fragebogen vorgestellt und abgefragt. Der angepasste *Manipulationscheck*

der Kundensteuerung wurde am Ende des jeweiligen Szenarios abgefragt. Verzerrungs- und Platzierungseffekte konnten vermieden werden, indem die Präsentation von Szenario B und D in der Reihenfolge variiert wurde.

6. Der Manipulationscheck für die experimentellen Bedingungen Geschäftsmodell und Kommunikation wurde im sechsten Teil durchgeführt. Zusätzlich wurde die abhängige Variable *ex post* Image abgefragt.
7. Im siebten Teil wurden die Probanden bezüglich ihrer *Flugerfahrung*, auf das Jahr 2006 bezogen, befragt.
8. Im letzten Teil des Online-Fragebogens konnten die Probanden ihre Kontaktdaten für den Wettbewerb angeben. Eine E-Mail-Adresse stand für Fragen zum Experiment zur Verfügung.

5.2.5 Pretests

Der erste Pretest wurde vor der Flughafen-Umfrage durchgeführt. Weil das zweite Experiment, die Online-Umfrage, Erkenntnisse der Flughafen-Umfrage berücksichtigte und weil bei web-unterstützten Experimenten aufgrund technischer Restriktionen neue Probleme auftreten können, wurde für den Online-Fragebogen ein zweiter Pretest durchgeführt (vgl. Frick et al., 2001).

Flughafen-Umfrage

Der Pretest wurde mit 30 Personen durchgeführt, zur Überprüfung der Manipulationen, der Struktur, der Verständlichkeit und der Dauer des schriftlichen Fragebogens. Zehn Personen wurden im Rahmen des Pretests beim Ausfüllen beobachtet und gebeten, ihre Gedanken mitzuteilen (*Think Aloud Method*; vgl. Mattila und Wirtz, 2006). Damit konnten Haltungen und Reaktionen der Probanden analysiert werden. Die Probanden bestanden aus Studenten der Universität St. Gallen, der ETH Zürich sowie weiteren Personen mit Flugerfahrung. Für den Pretest der englischen Version des Fragebogens wurden Probanden mit Muttersprache Englisch ausgewählt ($n = 8$).

Der Pretest zeigte, dass die Probanden über Dauer, Zweck und Vorgehen des Experiments verständlich instruiert wurden. Der Fragebogen konnte ohne Instruktion oder Unterstützung ausgefüllt werden. Für die Probanden waren die vorgestellten Szenarien und Situationen am Check-In sowie bei der Online-Buchung realitätsnah. Sie konnten sich gut in den hypothetischen Sachverhalt hineinversetzen. Die Fragen wurden von 29 Personen als verständlich bewertet. Das Antwortmuster bei den Probanden zeigte, dass eine natürliche Variabilität im Antwortverhalten besteht. Kein Proband war mit der Skalierung oder der Testdauer überfordert (im Durchschnitt 18.5 Minuten).

Kleinere Korrekturen am Fragebogen betrafen die Manipulation der Variablen *ex ante* Image und *Ticketpreis*. Sie mussten in einzelnen Szenarien angepasst werden, da Probanden beim Manipulationscheck Unsicherheiten formulierten. Weitere Korrekturen betrafen einzelne Fragestellungen, die

unter der Kundensteuerung (UV1), Faktorstufe „Anreiz“ unklar formuliert waren. Einzelne Probanden fühlten sich bei Erwähnung des experimentellen Versuchsaufbaus an eine Prüfungssituation erinnert. Die Probanden versuchten in diesem Fall herauszufinden, welche Manipulationen vorgenommen wurden, und passten das Antwortverhalten an. Dies führte zu einer Verzerrung im Antwortverhalten.

Ein erneuter Pretest mit neun Probanden bestätigte, dass die Anpassungen zu einem besseren Verständnis führten. Zudem fielen den Probanden die Manipulationen und der experimentelle Versuchsaufbau nicht mehr auf.

Online-Umfrage

Fünfzehn Personen, die mit verschiedenen Betriebssystemen und *Browser* arbeiten, wurde der Link zum Online-Fragebogen per E-Mail zugeschickt. Die Probanden stammten aus einer privaten E-Mail-Liste.

Die optische Veranschaulichung des Online-Buchungsvorgangs der Y-Air wurde von den Probanden als verständlich und realistisch bezeichnet: Die auf dem Internet simulierte Manipulation der Kommunikation mit einem deutlichen und einem versteckten Hinweis auf die Anreiz-Optionen und Tarifbedingungen bewährte sich (vgl. Abb. 5-4; Abb. 5-5). Sehr zufrieden waren die Probanden mit der kurzen Dauer der Befragung (10 Minuten) und den einfachen Rating-Skalen. Kein Proband hatte technische Probleme beim Ausfüllen oder Fragen zur Beantwortung des Fragebogens. Der Pretest zeigte, dass die neu eingeführte Manipulation des Geschäftsmodells deutlicher in der Ausgangssituation hervorgehoben werden musste. Da dieser Mangel nicht schwerwiegender Art war, wurde auf einen erneuten Pretest verzichtet.

5.3 Datenerhebung

Die praktische Umsetzung der Datenerhebung, die Gesamtstichprobengrösse und die Methode der Fragebogenauswertung wird in diesem Kapitel für beide Experimente (Flughafen- und Online-Umfrage) aufgeführt (Kap. 5.3.1; 5.3.2). Eine Übersicht zur Durchführung und statistischen Auswertung ist in Abbildung 5-7 dargestellt. Die Grundlagen zur statistischen Auswertung sind im Anhang aufgeführt (A.3).

5.3.1 Flughafen-Umfrage

Die Flughafen-Umfrage wurde an vier Tagen im Mai 2007 durchgeführt: am Wochenende vom 12. und 13. Mai sowie am Montag, 14. Mai, und Freitag, 25. Mai. Die Befragung an verschiedenen Werktagen sowie am Wochenende verminderte Verzerrungen bei der Stichprobenwahl zwischen Privat- und Berufspassagieren. Vier Studenten der Universität Zürich hatten Zugang zum Boarding-Bereich hinter der Sicherheitskontrolle am Flughafen Zürich (Gates A, B und E). Die Studenten

wurden in die Thematik des Fragebogens eingeführt, um allfällige Fragen der Passagiere beantworten zu können. Um eine möglichst gute Abbildung der Population „Passagiere“ in der Stichprobe zu erreichen, wurden die Studenten angewiesen, nicht länger als eine Stunde am selben Gate zu befragen. Diese Auswahlprozedur gewährleistete eine balancierte Durchmischung der Probanden nach Destination (Streckenlänge), Flugklasse, Airline und allgemeinen demografischen Variablen. Für die ausgewogene Zellenzuteilung des Versuchsplanes wurden pro Zelle gleich viele Fragebogenversionen gedruckt und randomisiert verteilt.

Abbildung 5-7: Durchführung und Auswertung der Experimente

	Durchführung	Statistische Auswertung
Experiment I (Schriftliche Umfrage)	<ul style="list-style-type: none"> – 12.-14. und 25. Mai 2007 am Flughafen Zürich – Population: Flugpassagiere – $N = 598$ (englische Version: 23%) – Auswahl: Randomisiert an Gates A, B, E Flughafen Zürich 	<ul style="list-style-type: none"> – Demografie: deskriptive Statistik – Manipulationscheck: <ul style="list-style-type: none"> – F-Test (ANOVA) – Hypothesentest: <ul style="list-style-type: none"> – t-Test (Mittelwerte) – F-Test (ANOVA) – Korrelationen (multiple Regression) – Effekt- und Teststärke – Modellprüfung <ul style="list-style-type: none"> – F-Test (ANOVA) – Punkt-biseriale Korrelationen
Experiment II (Online-Umfrage)	<ul style="list-style-type: none"> – 2. Juli - 10. August 2007 auf Webserver WEXTOR – Population: Internet-User – $N = 1'080$ (englische Version: 14%) – Auswahl: Internationale E-Mail-Listen, Newsgroups 	<ul style="list-style-type: none"> – vgl. oben (Experiment I) – Keine Modellprüfung mit F-Test und Korrelationen

Legende: N = Gesamtstichprobengröße, ANOVA = Analysis of Variance;

Quelle: Eigene Darstellung

Von den 636 am Flughafen eingesammelten Fragebogen wurden 38 als ungültig deklariert. Entweder waren weniger als 50% der Fragen beantwortet, die Fragebogen mit unseriösen Kommentaren versehen oder sie wiesen ein musterhaftes Antwortverhalten auf (z.B. bei allen Skalen tiefste Ausprägung gewählt). Auf sämtliche experimentellen Bedingungen verteilt gilt daher eine Gesamtstichprobengröße (N) von 598 für die Flughafen-Umfrage. Von den 598 Fragebogen waren 139 Stück (23%) in der englischen Version ausgefüllt worden. Die Anforderungen an die optimale Stichprobengröße wurden sowohl im schriftlichen als auch im Online-Experiment erfüllt. Die Aufteilung der Gesamtstichprobengröße auf die einzelnen Zellen der experimentellen Bedingungen ist in Tabelle 5-6 dargestellt:

Tabelle 5-6: Stichprobengröße (*n*) pro experimentelle Bedingung (Flughafen-Umfrage)

<i>n</i> pro Bedingung	UV1	UV2	UV3	UV4
	Kundensteuerung	Kommunikation	Ticketpreis	Image
Ausprägung „hoch“	323	306	293	311
Ausprägung „tief“	275	292	305	287

Quelle: Eigene Berechnung mit SPSS

5.3.2 Online-Umfrage

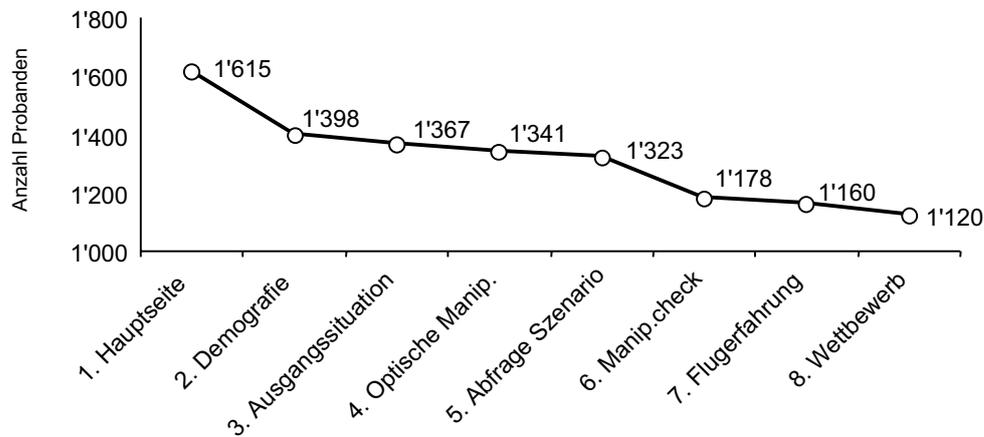
Das Sample der Online-Untersuchung setzte sich zusammen aus privaten E-Mail-Listen (397 Kontakte) sowie einer internationalen E-Mail-Datenbank mit 1'513 Versuchsteilnehmern der *Mobile Spam Study 2005* (Brodt und Heé, 2005). Um eine möglichst breite Verteilung der Stichprobe zu gewährleisten, wurde der Fragebogen in internationalen Gruppen-Foren auf www.xing.com aufgeschaltet sowie im Newsletter von www.umfragecenter.de erwähnt (6'730 Kontakte).

Die Verteilung erfolgte mit einer E-Mail auf deutsch und englisch, mit einem Linkverweis zum Online-Fragebogen in der jeweiligen Sprache. Der Versand der E-Mails wurde ab 2. Juli 2007 innerhalb von drei Tagen gestaffelt durchgeführt. Es wurde angenommen, dass sich viele Probanden während dieser Zeit in den Vorbereitungen zu einer Ferien- oder Flugreise befanden. Inhaltlich betrachtet, war damit ein leichter Zugang zum Fragebogen möglich.

Der Fragebogen blieb bis zum 10. August 2007 auf dem WEXTOR Server des Wirtschaftspsychologischen Instituts der Universität Zürich erreichbar. Total besuchten während dieser Zeit 1'615 Personen die erste Seite des Online-Fragebogens. 1'340 *Hits* betrafen die deutsche Version des Fragebogens, 275 *Hits* die englische Version (Doppelzählungen berücksichtigt). Von diesen total 1'615 *Hits* füllten schliesslich 1'120 Probanden den ganzen Fragebogen aus. Die Ausfallquote für beide Fragebogen beträgt 31%. Das *Drop-out-Profil* ist in Abbildung 5-8 für beide Fragebogen pro Seite der Online-Umfrage dargestellt. Durchschnittlich benötigten die Probanden 9.5 Minuten (Standardabweichung (*s*): 6.3, *Range*: 6 bis 18 Minuten) zum Ausfüllen des Fragebogens.

Von den 1'340 Personen, welche den deutschen Fragebogen besuchten, wurden 24 ausgeschlossen, weil sie nicht ernsthaft teilnehmen wollten oder keine Flugerfahrung hatten (*Seriousness Check*). Zehn Probanden beantworteten weniger als 50% der Fragen oder der Fragebogen wies ein musterhaftes Antwortverhalten auf. Schliesslich fielen 379 Personen aus der Stichprobe, weil sie das Experiment während des Durchgangs abgebrochen haben (149 Probanden bereits nach der Hauptseite). Die endgültige Stichprobengröße lag somit auf 927 Probanden für die deutsche Version des Fragebogens.

Von 275 Personen, welche die englische Version des Fragebogens besuchten, brachen 116 das Experiment während des Durchgangs ab (68 Probanden nach der Hauptseite). Mehrere Forschungs-

Abbildung 5-8: Ausfallquote web-gestützter Fragebogen

Legende: Manip. = Manipulation (der unabhängigen Variablen);

Quelle: Eigene Darstellung

arbeiten bestätigen, dass die durchschnittliche Ausfallquote bei Online-Umfragen höher ist als bei schriftlichen Umfragen (Janetzko et al., 2002; Birnbaum, 2000; Musch und Reips, 2000). Die Validität der vorliegenden Untersuchung ist daher nicht eingeschränkt. Drei Personen wurden aufgrund des *Seriousness Checks* von der Analyse ausgeschlossen. Drei Fragebogen wiesen musterhaftes Antwortverhalten auf oder mehr als 50% der Antworten fehlten. Für die Auswertung konnten somit 153 gültige, englische Fragebogen verwendet werden (14% der Gesamtstichprobe).

Tabelle 5-7: Stichprobengrösse (n) pro experimentelle Bedingung (Online-Umfrage)

n pro Bedingung (Online)	UV1	UV2	UV5
	Kundensteuerung	Kommunikation	Geschäftsmodell
Ausprägung „hoch“	559	539	-
Ausprägung „tief“	521	541	-
Ausprägung „Network Carrier“	-	-	270
Ausprägung „Low Cost Carrier“	-	-	248
Ausprägung „Charter Airline“	-	-	273
Ausprägung „Regional Airline“	-	-	289

Quelle: Eigene Berechnung mit SPSS

Die Gesamtstichprobengrösse N der Online-Umfrage, welche zur Auswertung zur Verfügung stand, betraf für das zweite Experiment 1'080 Probanden. Die Aufteilung dieser Probanden auf die experimentellen Bedingungen der Online-Umfrage ist in Tabelle 5-7 dargestellt.

5.4 Programme zur statistischen Auswertung

Die Daten des Flughafen- und des Online-Experiments wurden mit SPSS (Statistical Package for the Social Science) Version 13 für Mac OS X ausgewertet. SPSS (o.V., 2006c) ist ein modular aufgebautes Statistik-Analyse-System zum Verwalten und statistischen Auswerten von Datenmaterial. Folgende Funktionen von SPSS standen für die Datenauswertung zur Verfügung: Daten einlesen und bearbeiten, Q-Q (Quantil-Quantil) Plots sowie Interaktions-Diagramme erzeugen und statistische Verfahren berechnen (vgl. A.3). Die vorliegende Datenauswertung wurde mit der SPSS Programmiersprache durchgeführt, indem die benötigten SPSS Funktionen in ein Syntax-Fenster geladen und ausgeführt wurden. Dadurch sind die Reproduzierbarkeit der Ergebnisse und die Durchführung gleichartiger Analysen auf mehreren Datensätzen möglich.

Der optimale Stichprobenumfang, die Teststärke und das Differenzmass wurden mit G*Power 3.0.6 für Mac OS X berechnet (Faul et al., 2007a). Diese statistischen Auswertungen sind in SPSS nicht möglich oder schwierig zu berechnen. G*Power ist ein statistisches Analyseprogramm zur Auswertung von t -, F - und χ^2 -Tests. Zusätzlich ermöglicht G*Power die Berechnung der statistischen Power (Teststärke), das Differenzmass und die optimale Stichprobengrösse pro Zelle im Experiment (Faul et al., 2007b). Die Berechnung dieser Kennwerte erfolgte durch die manuelle Eingabe der in SPSS berechneten Werte (vgl. Kap. 5.1.6).

5.4.1 Flughafen-Umfrage

Die Auswertung der Fragebogen erfolgte in einem ersten Schritt durch die manuelle Eingabe der Antworten in die Software *Microsoft Excel*. Nachdem alle Daten erfasst waren, wurde die Excel-Datei in SPSS importiert. Die Auswertung erfolgte in einem zweiten Schritt in SPSS, basierend auf dem importierten Datenblatt. Die statistischen Methoden zur Datenanalyse werden in Abbildung 5-7 für beide Experimente beschrieben.

5.4.2 Online-Umfrage

Die Software WEXTOR ermöglichte den Download sowohl der Antwortdaten als auch des *Log-File*. Die Antwortdaten standen nach dem Download im Excel-Format zur Verfügung, das in SPSS importiert wurde. Das *Log-File* erlaubte die Analyse von Zugriffszeiten, Dauer der Befragung insgesamt und pro Seite, *Drop-out-Profil* sowie die IP-Adresse der Probanden (vgl. Reips und Stieger, 2004). Mit dem *Log-File* konnten von einer Person mehrfach ausgefüllte Fragebogen herausgefiltert werden und Statistiken zur Ausfallquote erstellt werden.

5.5 Zusammenfassung

In diesem Kapitel wurden die Grundlagen zur experimentellen Forschung vorgestellt und der Fragebogen für die Flughafen- und die Online-Befragung erarbeitet. Die Kombination eines schriftlichen Flughafen- und eines web-unterstützten (Labor-) Experiments im *strengen Between-Subject-Design* und die Durchführung einer Exploration (Pretest, Manipulationscheck) erhöhten die Güte der vorliegenden Untersuchung.

Fünf experimentelle Bedingungen aus dem Forschungsmodell wurden mit der Szenario-Methode sowie einer grafisch angepassten Internetseite manipuliert. Dabei wurde die Auswirkung der unabhängigen auf die abhängigen Variablen in verschiedenen Szenarien mit Likert-Skalen gemessen. Die Szenarien betrafen je nach Kundensteuerung *Anreiz-Optionen* bei der Online-Buchung oder *Bestrafungsrichtlinien* am Check-In.

Das schriftliche Experiment am Flughafen Zürich wurde ohne Versuchsleiter durchgeführt, das Online-Experiment durch Versand eines Links zum web-basierten Fragebogen. Die statistischen Vorgaben für die optimale Stichprobengröße wurden erfüllt: Pro Hypothese standen nach dem Manipulationscheck mindestens 198 Probanden zur Verfügung. Im Flughafen- sowie im Online-Experiment wurden die Vorgaben zur Gesamtstichprobengröße mit $N = 598$ und $N = 1'080$ erreicht. Die statistische Auswertung wurde mit SPSS und G*Power durchgeführt.

6 Forschungsergebnisse

Das zentrale Forschungsmodell wird nach erfolgter Datenerhebung getestet. Ziel dieses Kapitels ist es, als konfirmatorischer Teil der quantitativen Untersuchung die Ergebnisse der beiden Experimente darzulegen und zu interpretieren. In Kapitel 6.1 werden zuerst die Ergebnisse der deskriptiven Datenanalyse aufgeführt. Nachdem in Kapitel 6.2 das Vorgehen und die Grundlagen zur quantitativen Datenanalyse kurz beschrieben werden, folgen die Forschungsergebnisse.

Im ersten Experiment (Flughafen-Umfrage) wird das Modell mit den Hypothesen (h) 1 bis 7 getestet und der Einfluss der demografischen Variablen untersucht (Kap. 6.3). Neben dem Hypothesentest wird das Gesamtmodell auf Korrelationen und Interaktionen zwischen den UV analysiert (Kap. 6.4). Aufgrund von Erkenntnissen des ersten Experiments werden in Kapitel 6.4 Modifikationen am Forschungsmodell vorgenommen und damit der Online-Fragebogen angepasst. Darauf aufbauend wird das zweite, web-gestützte Experiment durchgeführt und anschliessend getestet (Kap. 6.5). Die empirischen und methodischen Grenzen sind in Kapitel 6.6 aufgeführt. Die Folgerungen und damit das überprüfte Modell werden in Kapitel 6.7 vorgestellt.

6.1 Deskriptive Datenanalyse

Die deskriptive Datenanalyse betrifft die demografischen Daten der Probanden und die Merkmale zum Flugverhalten. Sie wird in diesem Kapitel für das Flughafen-Experiment (Experiment I) und für die Online-Umfrage (Experiment II) einzeln aufgeführt.

6.1.1 Flughafen-Experiment

Demografische Merkmale der Probanden

Unter den 598 gültigen Fragebogen sind die demografischen Daten zum Alter näherungsweise normalverteilt, mit einem Durchschnittsalter von 37.7 Jahren (Standardabweichung: $s = 14.41$, *Range*: 15 bis 83 Jahre). Mit 46% weiblichen und 54% männlichen Teilnehmern halten sich geschlechtsspezifischen Merkmale die Waage. Die meistgenannte Ausbildung ist Universität / Hochschule (38%), gefolgt von Mittelschule / Gymnasium und Lehre / Berufsschule (beide 19%). Über die Hälfte (56%) der Probanden haben ihren Wohnsitz in der Schweiz. 15% der Probanden wohnen in Deutschland, 7% in den USA und 5% in Irland. Der Rest teilt sich auf in 12% „andere Europa“ und 5% „andere“. In Tabelle 6-1 sind die demografischen Merkmale „Alter“, „Geschlecht“, „Ausbildung“ und „Wohnsitz“ für beide Experimente aufgeführt. Die Merkmale „Alter“ und „Geschlecht“ der Stichprobe repräsentieren das Passagierprofil der *Airport Service Quality* Studie von Unique (o.V., 2006d).

Tabelle 6-1: Demografisches Profil (Flughafen- und Online-Umfrage)

		Flughafen-Umfrage N = 598		Online-Umfrage N = 1'080	
Alter	Durchschnittsalter	37.7	-	30.6	-
	max. Alter	83	-	68	-
	min. Alter	15	-	16	-
Geschlecht	männlich	316	54%	556	52%
	weiblich	273	46%	520	48%
Ausbildung	Universität / Hochschule	222	38%	616	57%
	Höhere Fachschule	55	9%	56	5%
	Meisterdiplom / Fachausweis	48	8%	35	3%
	Mittelschule / Gymnasium	112	19%	296	28%
	Lehre / Berufsschule	110	19%	55	5%
	Schulobligatorium	28	5%	8	1%
	andere	13	2%	14	1%
Wohnsitz	Schweiz	328	56%	480	45%
	Deutschland	91	15%	426	39%
	USA	41	7%	44	4%
	Irland (Online: Singapur)	27	5%	35	3%
	andere Europa	72	12%	54	5%
	andere	30	5%	39	4%

Quelle: Eigene Darstellung

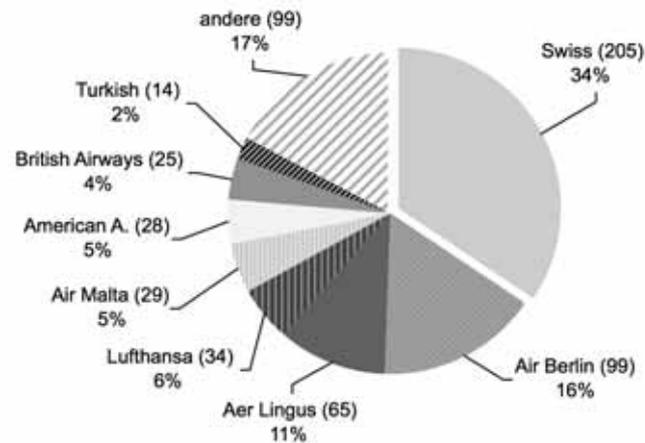
Demografische Merkmale zum Flugverhalten

Die befragten Passagiere flogen zum Zeitpunkt der Untersuchung zum grössten Teil innerhalb Europas (76%) oder nach Nord- / Südamerika (13%). In den Nahen Osten flogen 7%, nach Asien / Australien 4% der Befragten (N = 590). Dabei hatten 57% der Passagiere ihren Flug online gebucht. Erwartungsgemäss flogen vom Flughafen Zürich die meisten Probanden (N = 598) mit Swiss (34%), gefolgt von Air Berlin (16%), Aer Lingus (11%) und Lufthansa (6%). Insgesamt wurden 32 verschiedene Airlines genannt (vgl. Abb. 6-1). Der grösste Teil der Passagiere buchte in der *Economy Class* (92.5%), während 7% *Business Class* und nur 2 Probanden *First Class* flogen. Von 594 Probanden machten 36% bei einem Vielfliegerprogramm mit, wobei 41% auf dem „Einstiegsniveau“ und 46% „Vielflieger“ sind. 10% der Probanden sind „Gold Member“ und 3% nennen ihren Status „höchstes Niveau“. Diese (flug-) demografischen Merkmale der Stichprobe decken sich grösstenteils mit Auswertungen zu den Verkehrszahlen vom Flughafen Zürich (o.V., 2006d).

Im Durchschnitt wurden im Jahr 2006 11.1 Flüge ($s = 16.42$) für geschäftliche Zwecke und 3.7 ($s = 4.12$) für private Zwecke gebucht. 14% der Probanden sind im Jahr 2006 nicht geflogen. Die Vielflieger sind vor allem geschäftlich unterwegs: Von 246 Probanden, die im 2006 geschäftlich geflogen sind, haben 17% mehr als fünfzehn Flüge absolviert. Für private Zwecke sind nur 2% mehr als fünfzehn Mal geflogen. 68% von 480 Probanden sind im 2006 für private Zwecke ein bis

drei Mal geflogen (geschäftlich: 36%), 17% vier bis sechs Mal (geschäftlich: 18%) und 5% sieben bis neun Mal (geschäftlich: 12%).

Abbildung 6-1: Geflogene Airlines am Tag der Umfrage (Flughafen-Umfrage)



Legende: A. = Airlines

Quelle: Eigene Darstellung

6.1.2 Online-Experiment

Demografische Merkmale der Probanden

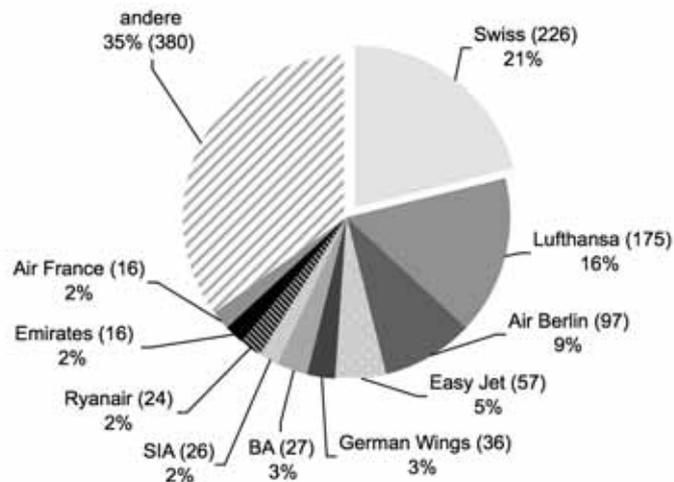
Die Online-Stichprobe weist eine ähnliche demografische Verteilung der Geschlechts- und Ausbildungsmerkmale auf wie die Flughafen-Stichprobe. 52% sind männliche Probanden und 48% sind weiblich ($N = 1'077$). Am meisten genannt wurde die Universität / Hochschule (57%) als abgeschlossene Ausbildung, gefolgt von Mittelschule / Gymnasium (28%) und Lehre / Berufsschule sowie höhere Fachschule (beide 5%). Das Durchschnittsalter ist mit 30.6 Jahren bei den 1'080 Internet-Probanden tiefer als in der Flughafen-Umfrage ($s = 11.31$, *Range*: 16 bis 70 Jahre). Von 1'072 antwortenden Probanden haben 71% ihren letzten Flug online gebucht. 11% der Probanden buchten ihren Flug nicht selber. Der Wohnsitz der Probanden ist ähnlich verteilt wie in der Flughafen-Umfrage. Den grössten Anteil machten Probanden mit Wohnsitz in der Schweiz aus (45%), gefolgt von Deutschland (39%), USA (4%) und Singapur (3%).

Demografische Merkmale zum Flugverhalten

Bei der Frage nach der Airline, mit der im Jahr 2006 meistens geflogen wurde, wurden 85 verschiedene Fluggesellschaften genannt (vgl. Abb. 6-2). Dies bestätigt die im Vergleich zur Flughafen-Umfrage breitere demografische Streuung der Online-Stichprobe. Die Reihenfolge der meist geflogenen Fluggesellschaften sieht ähnlich aus: Swiss (21%), Lufthansa (16%), Air Berlin (9%) und EasyJet (5%) sind die meist genannten Airlines ($N = 1'080$). Mit 92% *Economy-Class*- und 7%

Business- resp. 0.3% *First-Class*-Passagieren weist die Online-Umfrage eine fast identische Klassenstruktur wie die Stichprobe der schriftlichen Umfrage auf. Im Jahr 2006 flogen die Probanden im Durchschnitt 7.3 Mal (privat und geschäftlich, $s = 9.51$). Dies entspricht ebenfalls dem Durchschnitt der Flughafen-Untersuchung. Im Gegensatz zur Flughafen-Umfrage sind in der Stichprobe der Online-Umfrage weniger Probanden (8%) im Jahr 2006 nicht geflogen.

Abbildung 6-2: Meist geflogene Airlines im 2006 (Online-Umfrage)



Legende: BA = British Airways, SIA = Singapore Airlines

Quelle: Eigene Darstellung

6.2 Vorgehen und Methoden der statistischen Datenanalyse

Jede Hypothese wird in der vorliegenden Untersuchung anhand des folgenden Schemas strukturiert und geprüft (i. A. an Hirsig, 2002):

1. Formulierung der Hypothese (h)

(a) Inhaltlich-theoretische Formulierung der Hypothese

- i. Die Hypothese wird, basierend auf der inhaltlichen Formulierung aus Kapitel 4.2, modelltheoretisch formuliert.
- ii. Grundlage der Hypothesenformulierung sind der theoretische Input und die Erkenntnisse der qualitativen Empirie (vgl. Kap. 3.8).

(b) Statistische Formulierung der Hypothese (vgl. A.3)

- i. Der Nullhypothese (H_0), welche postuliert, dass beide Stichproben aus derselben Population stammen, wird die Alternativhypothese (H_1) gegenübergestellt (vgl. Kromrey, 2006).

- ii. Da eine irrtümliche Annahme der Nullhypothese in der vorliegenden Untersuchung weniger gravierend ist als eine irrtümliche Ablehnung, wird der Fehler erster Art (α) möglichst tief gehalten. Dabei wird der Fehler 2. Art (β) grösser und damit die Teststärke ($1 - \beta$) tendenziell tiefer (vgl. A.3.2; Cramer und Kamps, 2007; Schira, 2003).

2. Definition der Stichprobengrösse (n)

- (a) Die Stichprobengrösse wird in jeder Hypothese durch den Manipulationscheck der betreffenden unabhängigen Variable (UV) definiert. Zu berücksichtigen sind die Mindestanforderungen aus der Poweranalyse (vgl. Kap. 5.1.6).
- (b) Wenn die Hypothese eine unabhängige Variable betrifft, wird zuerst geprüft, ob es Unterschiede zwischen den Gruppen „Manipulationscheck funktioniert“, „Manipulationscheck nicht funktioniert“, „neutrale Antwort“ und „keine Antwort“ gibt. Bestehen im Antwortverhalten der Fairness-Beurteilung (AV3) signifikante Unterschiede zwischen diesen Gruppen, wird für den Hypothesentest jene Substichprobe gewählt, bei der die Manipulation funktioniert hat. Ergeben sich keine signifikanten Mittelwertsunterschiede zwischen den Gruppen des Manipulationschecks, können alle gültigen Fragebogen verwendet werden.
- (c) Betrifft die Hypothese ausschliesslich im Modell als abhängige Variablen erhobene Faktoren, werden alle gültigen Fragebogen verwendet ($N = 598$).

3. Hypothesen prüfen

(a) t -Test für Mittelwertsvergleiche

- i. Für den Vergleich der Mittelwerte (\bar{x}) der unabhängigen Stichproben wird pro Szenario (A bis D) mit dem t -Test geprüft. Jeder Mittelwert wird mit der dazugehörigen Standardabweichung (Stand.abw.) s angegeben.
- ii. Pro Szenario werden zwei Mittelwerte verglichen: Die erste Gruppe umfasst alle Probanden, welche die UV Faktorstufe „hoch“ (bei UV1: Anreiz) zugewiesen bekommen. In der zweiten Gruppe erfolgt die Manipulation mit der Faktorstufe „tief“ (bei UV1: Richtlinie). Die beiden Mittelwerte der Fairness-Beurteilung sind auf einer 7er Likert-Skala von 1 für „sehr unfair“ bis 7 für „sehr fair“ skaliert. Die ungerade Anzahl Kategorien ermöglicht eine Position „neutral“, welche mit 4 kodiert ist.
- iii. Die Nullhypothese wird abgelehnt, wenn die Überschreitungswahrscheinlichkeit (p) < 0.05 ist. Der Zufall kann für die Interpretation der Mittelwertsunterschiede mit einer Irrtumswahrscheinlichkeit kleiner 5% ausgeschlossen werden (vgl. A.3.2). In

den Sozialwissenschaften haben sich die Irrtumswahrscheinlichkeiten (Signifikanzniveaus) von 5%, 1% und 0.1% etabliert (vgl. Hamlin, 2005; Hussy und Jain, 2002; Hicks, 1993). Die Übertretungswahrscheinlichkeit (p) wird auch Signifikanz (Sig.) genannt (Duller, 2007). Kann die Nullhypothese nicht falsifiziert werden, handelt es sich um ein nicht signifikantes (n.s.) Ergebnis.

(b) Korrelationen

- i. Mit dem Korrelationskoeffizienten (r) wird die Wirkungsrichtung (gleichsinnig, gegensinnig) sowie die stochastische „Enge“ zwischen zwei Variablen beschrieben (Schnell et al., 1999). Er wird in der vorliegenden empirischen Untersuchung eingesetzt, wenn keine Gruppen mit Mittelwerten verglichen werden können (vgl. A.3.1).
- ii. Für die Beziehung zwischen UV und AV wird mit punkt-biserialen Korrelationen gerechnet, wenn es sich um eine Beziehung zwischen dichotom- und intervallskalierten Variablen handelt. Die punkt-biserielle Korrelation basiert auf der Pearson-Bravais-Korrelation und wird mit einer Produkt-Moment-Korrelation berechnet. Die Korrelationen zwischen AV werden mit der Produkt-Moment-Korrelation berechnet (Hirsig, 2002).
- iii. Bei 100 Freiheitsgraden (df) wird als Grenzwert für ein 5%-Signifikanzniveau $r = .20$ gefordert, bei 300 Freiheitsgraden $r = .11$ (vgl. A.3.1). Zur Hypothesenprüfung wird die Mindestgrösse für r je nach Anzahl Freiheitsgrade gemäss Hirsig (2002) bestimmt (vgl. Tab. A-2).

(c) Varianzanalyse (ANOVA)

- i. Für die Prüfung der Hypothesen 8 bis 10 im Online-Experiment wird mit SPSS eine Analysis of Variance (ANOVA) durchgeführt. Die Varianzanalyse ist die einschlägige Methode zur simultanen Prüfung der Auswirkungen mehrerer Stichprobenmittelwerte (vgl. A.3.3; Hussy und Jain, 2002; Hicks, 1993).
- ii. Für die Varianzanalyse müssen zwei Voraussetzungen erfüllt sein: die AV müssen normalverteilt sowie intervallskaliert sein und für alle Stichproben werden homogene Varianzen vorausgesetzt (vgl. Bamberg et al., 2007; Hicks, 1993). Mit SPSS wird diese Voraussetzung jeweils mit dem Levene-Test überprüft.
- iii. Das Vorgehen zur Berechnung der p -Werte und Schätzung der Varianzen mit dem F -Test mit einer ANOVA wird im Anhang beschrieben (A.3.3).

4. Beurteilung des nachgewiesenen Effekts (Effektstärke) und der Teststärke (*Power*)

- (a) Anhand der Berechnung der Post-hoc-Effektgrösse wird die praktische Bedeutung der Ergebnisse verifiziert, unabhängig von der Stichprobengrösse (vgl. A.3.4). Das Differenzmass (d) beträgt 0.2 für einen schwachen, $d = 0.5$ für einen mittleren und $d = 0.8$ für einen starken Effekt (Cohen, 1988).

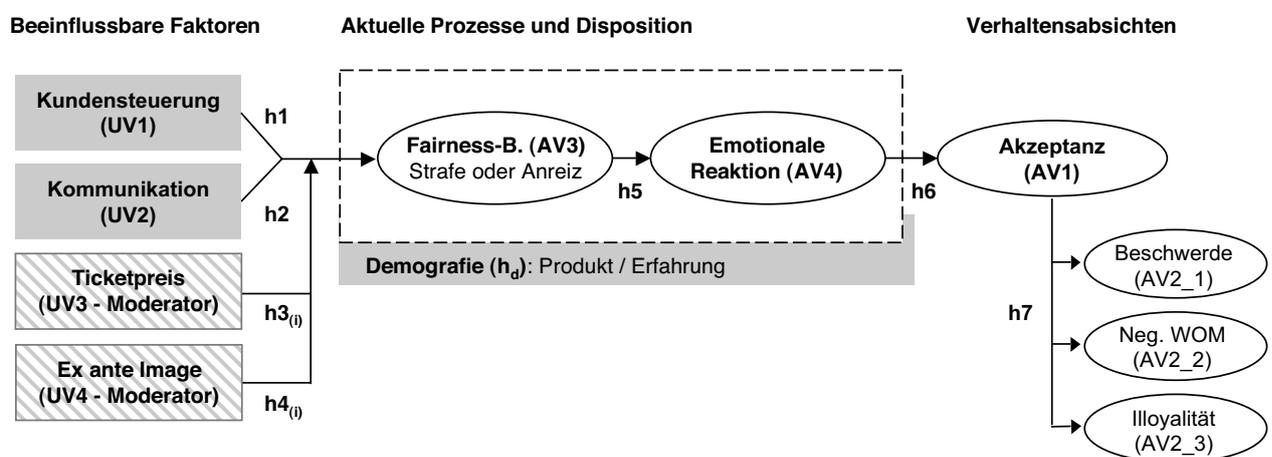
- (b) Die Effektgrösse ermöglicht zudem den Vergleich zwischen den Szenarien A bis D, was mit dem Wert der Übertretungswahrscheinlichkeit (p) nicht möglich ist (Faul et al., 2007b).
- (c) Mit der Post-hoc-Berechnung der Teststärke ($1 - \beta$) wird die Wahrscheinlichkeit angegeben, mit der eine richtige Alternativhypothese angenommen und die Nullhypothese korrekterweise abgelehnt wurde (vgl. A.3.2; Kirk, 1995).

5. Inhaltliche Interpretation bezüglich der formulierten Hypothese und Übersicht zu den statistischen Kennwerten der untersuchten Variablen.

6.3 Hypothesentest Experiment I

In den folgenden Kapiteln (6.3.1 bis 6.3.7) werden die Ergebnisse des Hypothesentests des Flughafen-Experiments (Experiment I) pro Hypothese aufgeführt und analysiert. Die Nebenhypothesen betreffen den Einfluss von demografischen Variablen (Klassenzugehörigkeit und Flugerfahrung) auf das Akzeptanzniveau und die negativen Verhaltensabsichten. Sie werden in Kapitel 6.3.8 diskutiert.

Abbildung 6-3: A-priori-Forschungsmodell Experiment I



Quelle: Eigene Darstellung

In Abbildung 6-3 ist das Forschungsmodell für das erste Experiment mit den untersuchten Faktoren sowie negativen Verhaltensabsichten abgebildet. Es basiert auf dem zentralen A-priori-Forschungsmodell, unter Berücksichtigung der relevanten, zu untersuchenden Variablen (vgl. Kap. 4.2). Die Wirkbeziehung zwischen den Variablen werden als Hypothesen (h) 1 bis 7 dargestellt. Der Einfluss auf die Hauptbeziehung, ausgehend von der *Kundensteuerung* und deren *Kommunikation* auf die *Fairness-Beurteilung*, wird von den Moderator-Variablen *Ticketpreis* und *ex ante Image* beeinflusst. Die im Modell postulierten negativen Verhaltensabsichten werden einzeln aufgeführt.

6.3.1 Kundensteuerung und Fairness-Beurteilung (h1)

1. Formulierung Hypothese 1

h1 modelltheoretisch formuliert: Wenn die Faktorstufe von UV1 (Kundensteuerung) tief (Richtlinie) ist, dann folgt eine tiefe Ausprägung der AV3, Fairness-Beurteilung (Beurteilung als unfair).

2. Stichprobengröße

Da die Kundensteuerung dichotom skaliert ist („Anreiz“ vs. „Richtlinie“) und mit zwei intervallskalierten Werten in Beziehung gesetzt wird, erfolgt der Manipulationscheck in SPSS durch Überprüfung mittels der punkt-biserialen Korrelation (vgl. A.3.1). Demnach müssen für eine erfolgreiche Manipulation über die gesamte Stichprobe hinweg zwei Bedingungen erfüllt sein:

1. Für den Manipulationscheck „Zusatznutzen“ muss eine positive (gleichsinnige) Korrelation mit der Kundensteuerung bestehen. Wenn bspw. die Kundensteuerung als „Anreiz“ formuliert ist, dann hat der Proband ebenfalls eine hohe Bewertung auf der Skala „Zusatznutzen“ angegeben und somit die erste Bedingung für den Manipulationscheck erfüllt (vgl. Kap. 5.2.3).
2. Die zweite Bedingung betrifft den Manipulationscheck „negative Konsequenz“. Die Korrelation zwischen der Kundensteuerung und „negativer Konsequenz“ muss negativ (gegensinnig) sein. Wenn bspw. die Kundensteuerung als Richtlinie formuliert ist, dann hat der Proband für einen „bestandenen“ Manipulationscheck einen höheren Wert auf der Skala gewählt.

Eine Berechnung der Korrelationen ist nur über die gesamte Stichprobe ($N = 598$) möglich: Da eine Überprüfung einzelner Probanden aus statistischen Gründen nicht möglich ist, wird pro Szenario die Erfüllung beider Bedingungen gefordert.

In jedem Szenario werden die beiden Anforderungen an die Korrelation erfüllt: in Szenario A auf dem $p < 1\%$ Niveau, in B auf dem $p < 5\%$ Niveau und in C sowie D auf dem Niveau $p < 0.1\%$. Bei der Interpretation der Korrelationen in C oder D kann somit der Zufall mit einer Irrtumswahrscheinlichkeit kleiner als 0.1% ausgeschlossen werden. Die einzelnen r -, p - und n -Werte sind in Tabelle 6-2 aufgeführt.

In der Gesamtbetrachtung hat die Manipulation *Kundensteuerung* funktioniert und alle Fragebogen können für h1 ausgewertet werden. Für die Faktorstufe „Richtlinie“ beträgt die Stichprobengröße somit $n = 275$ und für die Faktorstufe „Anreiz“ $n = 323$.

3. Hypothesenprüfung (Mittelwertsunterschiede)

Zur Hypothesenprüfung werden die Mittelwerte der abhängigen Variable *Fairness-Beurteilung* unter den verschiedenen Faktorstufen der Kundensteuerung ermittelt. Mit dem t -Test werden anschließend die Mittelwerte der Kundensteuerung „Richtlinie“ und der Kundensteuerung „Anreiz“ pro Szenario A bis D miteinander verglichen.

Für die Werte n , \bar{x} , s , df , t , p und d wird für jedes Szenario A bis D auf Tabelle 6-2 verwiesen. Zur Ergänzung wird die punkt-biseriale Korrelation (P.b.Korr.) zwischen der UV *Kundensteuerung* und der AV *Fairness-Beurteilung* angegeben.

Tabelle 6-2: Auswirkung von Kundensteuerung auf Fairness-Beurteilung

h1: UV1 → AV3		Kundensteuerung		t-Test (2-seitig)			Effekt	P.b.Korr.
		Richtlinie $n = 275$	Anreiz $n = 323$	df	t	Sig.	d	r Sum: .57
A	Mittelwert	3.15	3.66	542	-4.50	.00	0.37	.18
	Stand.abw.	1.47	1.28					
B	Mittelwert	3.36	4.20	558	-7.43	.00	0.60	.29
	Stand.abw.	1.44	1.38					
C	Mittelwert	3.14	4.22	592	-9.85	.00	0.82	.38
	Stand.abw.	1.30	1.35					
D	Mittelwert	2.43	4.79	592	-23.54	.00	1.94	.67
	Stand.abw.	1.23	1.20					

Legende: A = Übergepäck, B = Lounge-Zugang / VIP Boarding, C = flexibles Rückreisedatum, D = Meilen-Upgrade;

Quelle: Eigene Berechnung mit SPSS und G*Power

Die beobachteten Unterschiede der Mittelwerte sind in allen Szenarien mit einer Irrtumswahrscheinlichkeit von $p < 0.1\%$ nicht zufällig entstanden. In allen Szenarien A bis D wird die Nullhypothese H_0 falsifiziert und die Alternativhypothese H_1 angenommen.

4. Effektgrösse und Teststärke

Der nachgewiesene Effekt der Mittelwertsunterschiede ist gemäss der Einteilung von Cohen (1988) in Szenario A und B mittelstark ausgeprägt ($d(A) = 0.37$, $d(B) = 0.60$). Mit einer Effektgrösse von 0.82 in C und 1.94 in D sind die nachgewiesenen Effekte stark ausgeprägt. Wird die Effektgrösse über alle Szenarien (A + B + C + D) hinweg gerechnet, resultiert ein starker Effekt von $d = 1.39$.

Aufgrund des hohen Differenzmasses, des Signifikanzniveaus und der Stichprobengrösse in den beiden Szenarien C und D folgt eine hohe Teststärke von $(1 - \beta) = 1$. In Szenario A und C ist die Teststärke mit $(1 - \beta) = .72$ und $.99$ ebenfalls hoch.

5. Interpretation und Übersicht (h1)

Für jedes Szenario von A (Übergepäck) bis D (Meilen-Upgrade) wird die Nullhypothese falsifiziert und die Alternativhypothese mit der Irrtumswahrscheinlichkeit $p < 0.1\%$ angenommen. Die Wahrscheinlichkeit, dass H_0 korrekterweise abgelehnt wird, ist hoch. Für die Probanden besteht ein signifikanter Unterschied, ob derselbe Sachverhalt als Richtlinie mit Sanktion oder Anreiz gestaltet wird. Die hohe Teststärke bestätigt dieses Ergebnis (in B, C, D). Kunden stufen dasselbe Szenario als „fairer“ ein, wenn es als Anreiz und nicht als Richtlinie gestaltet wird. Am ausgeprägtesten

ist diese Wahrnehmung gemäss Differenzwert in Szenario D (Meilen-Upgrade). Die ausgeprägte gleichsinnige Korrelation zwischen Kundensteuerungskonzept und Fairnesswahrnehmung bestätigt den Zusammenhang über alle Szenarien (A + B + C + D) hinweg.

6.3.2 Kommunikation und Fairness-Beurteilung (h2)

1. Formulierung Hypothese 2

h2 modelltheoretisch formuliert: Wenn die Faktorstufe von UV2 hoch ist (deutliche Kommunikation), dann folgt eine hohe Ausprägung der AV3 (Beurteilung als fair).

2. Stichprobengrösse

Für den Manipulationscheck *Kommunikation* werden zwei Skalen verwendet (vgl. Kap. 5.2.3). Eine erfolgreiche Manipulation der Kommunikation bedeutet, dass der Proband unter der Faktorstufe Kommunikation „hoch“ beim ersten Manipulationscheck „klar“ und beim zweiten „fair“ gewählt hat.

Die Auswertung der Stichprobe bezüglich Manipulationscheck *Kommunikation* ergibt folgende Gruppen: „Manipulation bei beiden Fragen funktioniert“ $n = 199$, „Manipulation bei beiden Fragen nicht funktioniert“ $n = 98$, „neutrale oder unklare Antwort“ $n = 270$ und „keine Antwort“ $n = 12$ (*Missings*: 19).

Im Gegensatz zu h1 werden zur Überprüfung der Manipulation der *Kommunikation* entscheidungsstatistische Prüfverfahren verwendet, da Gruppen mit Mittelwerten verglichen werden (ANOVA). Die Nullhypothese lautet: „Zwischen den verschiedenen Gruppen des Manipulationschecks UV2 gibt es keine Unterschiede bezüglich der (Fairness-) Mittelwerte.“

Mit der Varianzanalyse berechnet, wird die Nullhypothese in allen Szenarien abgelehnt: in Szenario A ($F = 3.01$, $df = 570$, $n = 571$, $p = .03$), B ($F = 3.29$, $df = 571$, $n = 572$, $p = .02$), C ($F = 4.20$, $df = 570$, $n = 571$, $p = .01$) und D ($F = 4.11$, $df = 569$, $n = 570$, $p = .01$).

Das Antwortverhalten unterscheidet sich, je nachdem ob der Manipulationscheck funktioniert hat oder nicht (resp. ob er neutral ausfällt oder keine Antwort gegeben wurde). Für den Hypothesentest h2 wird die Stichprobe Manipulationscheck UV2 „funktioniert“ verwendet: $n = 199$ für beide Faktorstufen.

3. Hypothesenprüfung (Mittelwertsunterschiede)

Zur Hypothesenprüfung werden die beiden Mittelwerte der AV *Fairness-Beurteilung* unter den Faktorstufen Kommunikation „unklar“ sowie „klar“ miteinander verglichen (t -Test pro Szenario A bis D). Die Werte n , \bar{x} , s , df , t , p und d sind für jedes Szenario A bis D einzeln in Tabelle 6-3 aufgeführt. Zur Ergänzung wird die punkt-biseriale Korrelation (P.b.Korr.) zwischen der UV *Kommunikation* und der AV *Fairness-Beurteilung* in Tabelle 6-3 angegeben.

Tabelle 6-3: Auswirkung von Kommunikation auf Fairness-Beurteilung

h2: UV2 → AV3		Kommunikation		t-Test (2-seitig)			Effekt	P.b.Korr.
		unklar n = 99	klar n = 100	df	t	Sig.	d	r Total: .33
A	Mittelwert	2.98	3.26	159	-2.90	.00	0.46	.22
	Stand.abw.	1.30	1.46					
B	Mittelwert	3.56	4.39	152	-3.90	.00	0.61	.28
	Stand.abw.	1.51	1.18					
C	Mittelwert	3.52	4.10	159	-2.45	.02	0.40	.19
	Stand.abw.	1.50	1.38					
D	Mittelwert	3.22	4.06	161	-3.10	.00	0.50	.24
	Stand.abw.	1.72	1.65					

Legende: Sig. „.00“ ist bei allen Hypothesen gerundet, weshalb verschiedene Irrtumswahrscheinlichkeiten resultieren.

Quelle: Eigene Berechnung mit SPSS und G*Power

Die Nullhypothese wird für h2 in allen Szenarien falsifiziert (A und D: $p < 1\%$ B: $p < 0.1\%$, C: $p < 5\%$). Die Annahme der Alternativhypothese bedeutet, dass bei einer verständlichen und klaren Kommunikation die Fairness-Beurteilung höher liegt als bei einer mangelhaften Kommunikation des Sachverhalts.

4. Effektgrösse und Teststärke

Bei der Manipulation des Faktors *Kommunikation* wird in allen Szenarien ausser C ein mittlerer Effekt nachgewiesen. Die Effektgrösse wird basierend auf den Mittelwertsunterschieden der Fairness-Beurteilung berechnet. Der Effekt ist mit $d = 0.61$ in Szenario B am stärksten, gefolgt von $d = 0.50$ in Szenario D. Das Szenario C weist mit einem Differenzmass von $d = 0.40$ einen Wert auf, der knapp unter der mittleren Stärke $d = 0.50$ liegt.

Die Post-hoc-Teststärke ($1 - \beta$) bewegt sich für alle Szenarien zwischen .43 (A) und .51 (D).

5. Interpretation und Übersicht (h2)

In allen Szenarien wird nachgewiesen, dass eine verständliche, klare und faire Kommunikation dazu beiträgt, dass Kunden einen Zwischenfall als fairer beurteilen. Am ausgeprägtesten ist diese Feststellung gemäss Effektgrösse in Szenario B (Lounge-Zugang) und D (Meilen-Upgrade). Die Nullhypothese wird in allen Szenarien falsifiziert.

6.3.3 Ticketpreis und Fairness-Beurteilung (h3)

1. Formulierung Hypothese 3

h3 modelltheoretisch formuliert: Wenn die Faktorstufe von UV3 (Ticketpreis) hoch ist, dann folgt eine hohe Ausprägung der AV3 (Beurteilung als fair).

2. Stichprobengröße

Die Manipulation des Ticketpreises (UV3) funktioniert bei 329 Probanden, bei 75 Probanden ist sie nicht erfolgreich, 170 Probanden antworten neutral und 3 Probanden möchten nicht antworten (*Missings*: 21). Die Nullhypothese zum Manipulationscheck *Ticketpreis* lautet analog zur Hypothese 2 (vgl. Kap. 6.3.2): „Zwischen den verschiedenen Gruppen des Manipulationschecks UV3 gibt es keine Unterschiede bezüglich der (Fairness-) Mittelwerte.“

Die Berechnung mit der ANOVA zeigt, dass die Nullhypothese in Szenario A und D mit einem p -Wert von .89 ($F = 0.22$, $df = 576$, $n = 577$) sowie $p = .67$ ($F = 0.52$, $df = 575$, $n = 576$) nicht falsifiziert wird. Es besteht kein signifikanter Unterschied der Fairness-Beurteilung zwischen den Gruppen „Manipulationscheck bestanden“, „nicht bestanden“, „neutral“ und „keine Antwort“. In Szenario B ($F = 2.60$, $df = 577$, $n = 578$, $p = .05$) und C ($F = 3.48$, $df = 575$, $n = 576$, $p = .02$) hingegen wird die Nullhypothese mit $p < 5\%$ abgelehnt. Daher wird zur Analyse von h3 für alle Szenarien die Stichprobe „Manipulationscheck bestanden“ verwendet.

Unter der Faktorstufe Ticketpreis „tief“ werden $n = 182$ Probanden untersucht, unter der Faktorstufe „hoch“ hat der Manipulationscheck *Ticketpreis* bei 147 Probanden funktioniert.

3. Hypothesenprüfung (Mittelwertsunterschiede)

Zur Hypothesenprüfung werden die beiden Mittelwerte der AV *Fairness-Beurteilung* unter den Faktorstufen Ticketpreis „hoch“ sowie „tief“ miteinander verglichen (t -Test pro Szenario A bis D). Die Werte n , \bar{x} , s , df , t , p und d sind für jedes Szenario A bis D einzeln in Tabelle 6-4 aufgeführt.

Tabelle 6-4: Auswirkung vom Ticketpreis auf Fairness-Beurteilung

h3: UV3 → AV3		Ticketpreis		t-Test (2-seitig)			Effekt	P.b.Korr.
		tief $n = 182$	hoch $n = 147$	df	t	Sig.		
A	Mittelwert	3.42	3.55	327	-0.82	.41	0.09	.05
	Stand.abw.	1.44	1.36					
B	Mittelwert	3.84	4.14	327	-1.92	.06	0.21	.11
	Stand.abw.	1.42	1.47					
C	Mittelwert	3.71	3.95	326	-1.48	.14	0.17	.08
	Stand.abw.	1.44	1.45					
D	Mittelwert	3.75	3.71	296	0.17	.86	0.02	-.01
	Stand.abw.	1.65	1.84					

Quelle: Eigene Berechnung mit SPSS und G*Power

In den Szenarien A bis D wird die Nullhypothese nicht falsifiziert. Zur Überprüfung der postulierten Wirkungsrichtung der Korrelation dient die punkt-biseriale Korrelation (P.b.Korr.) zwischen *Ticketpreis* und der Fairness-Beurteilung (vgl. Tab. 6-4).

4. Effektgrösse und Teststärke

In Szenario B und C besteht ein schwacher Post-hoc-Effekt bei den Mittelwertsunterschieden der Fairness-Beurteilung mit Differenzwerten von $d = 0.21$ (B) und $d = 0.17$ (C). Die Effektgrössen in den Szenarien A und D sind mit Werten kleiner als 0.10 schwach ausgeprägt. Die Mittelwerte der Fairness-Beurteilung (vgl. Tab. 6-4) in Szenario D sind bspw. für die Faktorstufe „tief“ $\bar{x} = 3.75$ ($s = 1.65$) und für Faktorstufe „hoch“ $\bar{x} = 3.71$ ($s = 1.84$). Dies führt zu einem geringen Differenzwert d von 0.02. Für die Summe aller Szenarien (A + B + C + D) gilt ein schwacher Effekt von $d = 0.19$.

Die Annahme der Nullhypothese ist mit einem β -Fehler von 14% am deutlichsten in Szenario D. Die Teststärke weist für die Szenarien A bis C Werte zwischen $(1 - \beta) = .51$ und $(1 - \beta) = .56$ auf. Mit einer hohen Wahrscheinlichkeit wurde die Nullhypothese fälschlicherweise angenommen.

5. Interpretation und Übersicht (h3)

Die Hypothese wird falsifiziert, es besteht nur ein schwacher Einfluss des Ticketpreises auf die Fairness-Beurteilung. Die Probanden in den Szenarien B (Lounge-Zugang) und C (Rückreisedatum) neigen zwar zu einer gleichsinnigen Berücksichtigung des Ticketpreises. Dieser Zusammenhang ist jedoch schwach ausgeprägt, was die Effektgrösse betrifft. In Szenario A (Übergepäck) und D (Meilen-Upgrade) besteht praktisch kein bedeutungsvoller Zusammenhang zwischen den verschiedenen Ticketpreisen. Da eine tiefe Teststärke vorliegt, ist weitere Forschung bezüglich der Ablehnung von h3 indiziert. Der Interaktionseffekt vom Ticketpreis wird in Kapitel 6.4.2 überprüft.

6.3.4 Ex ante Image und Fairness-Beurteilung (h4)

1. Formulierung Hypothese 4

h4 modelltheoretisch formuliert: Wenn die Faktorstufe von UV4 (*ex ante Image*) hoch ist, dann folgt eine hohe Ausprägung der AV3 (Beurteilung als fair).

2. Stichprobengrösse

Die deskriptive Statistik (ANOVA) bestätigt für den Manipulationscheck des *ex ante Image* ähnliche Werte wie beim *Ticketpreis*: Bei 292 Probanden hat die Manipulation funktioniert, bei 100 Probanden hat sie nicht funktioniert, 186 Probanden haben „neutral“ gewählt und 3 Probanden wählten „keine Antwort“ (*Missings*: 17).

Ausser in Szenario B ($F = 1.69$, $df = 581$, $n = 582$, $p = .17$) unterscheidet sich das Antwortverhalten in allen Szenarien. Für das Antwortverhalten werden die Mittelwerte der Fairness-Beurteilung der verschiedenen „Manipulationsgruppen“ miteinander verglichen. In Szenario A ($F = 5.90$, $df = 580$, $n = 581$, $p = .01$) und D ($F = 4.32$, $df = 579$, $n = 580$, $p = .01$) unterscheidet sich das Antwortverhalten auf dem Signifikanzniveau $p < 1\%$, in Szenario C ($F = 3.25$, $df = 579$, $n = 580$, $p = .02$) auf dem Signifikanzniveau $p < 5\%$. Daher werden zur Prüfung von h4 nur die Stichproben verwendet, welche den Manipulationscheck bestanden haben.

3. Hypothesenprüfung (Mittelwertsunterschiede)

Der *t*-Test für Mittelwertsgleichheit bestätigt, was die Gruppenstatistik vermuten lässt: Die Nullhypothese wird in den Szenarien B, C und D mit einer geringen Irrtumswahrscheinlichkeit von $p < 0.1\%$ falsifiziert. Für Szenario A wird die Nullhypothese auf dem $p < 1\%$ Niveau falsifiziert.

Verglichen werden die beiden Mittelwerte der AV *Fairness-Beurteilung* unter den Faktorstufen *ex ante Image* „hoch“ sowie „tief“ (*t*-Test pro Szenario A bis D). Die zur Hypothesenprüfung benötigten Werte n , \bar{x} , s , df , t , p , d (und ergänzend r) sind für jedes Szenario A bis D einzeln in Tabelle 6-5 aufgeführt.

Tabelle 6-5: Auswirkung des (*ex ante*) Image auf Fairness-Beurteilung

h4: UV4 → AV3		<i>Ex ante Image</i>		<i>t</i> -Test (2-seitig)			Effekt	P.b.Korr.
		tief $n = 176$	hoch $n = 116$	df	t	Sig.	d	r Total: .40
A	Mittelwert	3.30	3.74	290	-2.73	.01	0.32	.16
	Stand.abw.	1.33	1.42					
B	Mittelwert	3.49	4.56	288	-6.38	.00	0.76	.35
	Stand.abw.	1.38	1.43					
C	Mittelwert	3.44	4.25	289	-4.79	.00	0.57	.27
	Stand.abw.	1.38	1.48					
D	Mittelwert	3.15	4.35	289	-6.07	.00	0.73	.34
	Stand.abw.	1.69	1.59					

Quelle: Eigene Berechnung mit SPSS und G*Power

Zwischen dem *ex ante Image* und der *Fairness-Beurteilung* besteht ein Zusammenhang, der mit einer tiefen Wahrscheinlichkeit zufällig entstanden ist.

4. Effektgröße und Teststärke

Die Mittelwertsunterschiede der Fairness-Beurteilung führen, über alle Szenarien hinweg betrachtet (A + B + C + D), zu einer starken Effektgröße $d = 0.87$. Ein mittlerer Effekt wird in Szenario C mit $d = 0.57$ erreicht, während in Szenario A ein schwacher Effekt mit $d = 0.32$ besteht. Stark ausgeprägt ist hingegen der Differenzwert in Szenario B ($d = 0.76$) und in Szenario D ($d = 0.73$). In Szenario D erreicht die Faktorstufe *ex ante Image* „tief“ einen Durchschnittswert von $\bar{x}_1 = 3.15$ gegenüber $\bar{x}_2 = 4.35$ bei Image „hoch“. In Szenario B sind die \bar{x} -Werte 3.49 für die tiefe Faktorstufe resp. 4.56 für Image „hoch“, was zum starken Effekt führt (vgl. Tab. 6-5).

Die Teststärke für die Szenarien B (.99), C (.78) und D (.98) ist hoch. In Szenario A wird eine Teststärke von $(1 - \beta) = .48$ erreicht.

5. Interpretation und Übersicht (h4)

Es spielt bei einem Ereignis eine Rolle, welches Image eine Airline beim Kunden vor der Flugreise hat. Ist das *ex ante Image* hoch, wird ein negativer Sachverhalt milder, das heisst „fairer“ beurteilt (Auswirkung auf Fairness-Beurteilung). Airlines, deren Image tief ist, lösen bei Kunden mit hoher Wahrscheinlichkeit negative Fairness-Beurteilungen aus. Diese Aussage gilt vor allem in den Szenarien B (Lounge-Zugang), C (Rückreisedatum) und D (Meilen-Upgrade). Im Übergepäck-Szenario (A) spielt der Faktor *ex ante Image* hingegen nur eine schwache Rolle.

Weil der Zusammenhang unter h4 über alle experimentellen Bedingungen hinweg untersucht wird (d.h. für alle Faktorstufen der unabhängigen Variablen), wird die Aussagekraft der Interpretation gestärkt. Trotz einer ausgeprägten Teststärke besteht für die Variable *ex ante Image* weiterer Forschungsbedarf, da Image ein multidimensionales Konstrukt und somit anfällig ist für Fehler in der Operationalisierung. Der Interaktionseffekt vom Image wird in Kapitel 6.4.2 überprüft.

6.3.5 Fairness-Beurteilung und affektive Reaktion (h5)

1. Formulierung Hypothese 5

h5 modelltheoretisch formuliert: Wenn AV3 (Fairness-Beurteilung) tief ausgeprägt ist, dann folgt eine tiefe Ausprägung der AV4 (Verärgerung als emotionale Reaktion).

2. Stichprobengrösse

Mit h5 wird das Verhältnis zwischen zwei AV, der Fairness-Beurteilung und der emotionalen Reaktion, untersucht. Die Stichprobengrösse wird nicht durch die Manipulation beeinflusst, da keine experimentelle Bedingung unter h5 getestet wird. Daher werden alle Fragebogen in die Analyse für h5 mit einbezogen ($N = 598$).

3. Hypothesenprüfung (Korrelationen)

Wird die Stichprobe unabhängig vom Szenario untersucht, beträgt die Korrelation (A + B + C + D) $r = .87$ ($n = 598$, $p = .00$). In jedem einzelnen Szenario ist die Korrelation zwischen der Fairness-Beurteilung und der emotionalen Reaktion mit $r \geq .72$ stark gleichsinnig ausgeprägt und weist über 100 Freiheitsgrade auf. Für jedes Szenario gilt ein Signifikanzniveau von $p < 0.01\%$. Am deutlichsten ist das Ergebnis in Szenario D mit einem r -Wert von .89 (vgl. Tab. 6-6).

4. Effektgrösse und Teststärke

Jedes Szenario weist einen starken Effekt auf ($d_{A-D} \geq 2.06$) und eine hohe Korrelation mit $r_{Sum} = .87$ ($n = 598$, $p = .00$). Die Teststärke ist in allen Szenarien $(1 - \beta) = 1$.

Tabelle 6-6: Zusammenhang zwischen Fairness-Beurteilung und emotionaler Reaktion

h5: AV3 → AV4		Korrelationen ($n = 598$)		
Szenario	r	Sig. (2-seitig)	d	
A. Übergepäck	.72	.00	2.06	
B. Lounge / VIP	.78	.00	2.53	
C. Rückreisedatum	.82	.00	2.86	
D. Meilen-Upgrade	.89	.00	3.90	
Total (Szenario unabhängig)	.87	.00	3.45	

Quelle: Eigene Berechnung mit SPSS und G*Power

5. Interpretation und Übersicht (h5)

Das Konstrukt *emotionale Reaktion* ist mit der Fairness-Beurteilung hoch korreliert. Die Konstrukte sind sehr ähnlich: Ein Proband mit einer hohen Fairness-Beurteilung wird in fast allen Fällen eine positive emotionale Reaktion aufweisen. Diese Aussage ist auf einem Signifikanzniveau von $p < 0.01\%$ ($r = .87$, $n = 598$) und einer Teststärke von 1 gemacht (vgl. Tab. 6-6). Mit der Messung der Fairness-Beurteilung kann somit die emotionale Reaktion vorhergesagt werden. Im zweiten Experiment wird aus diesem Grund auf die Messung der emotionalen Reaktion verzichtet.

6.3.6 Emotionale Reaktion und Akzeptanzniveau (h6)

1. Formulierung Hypothese 6

h6 modelltheoretisch formuliert: Wenn die Ausprägung von AV4 tief ist (Verärgerung als emotionale Reaktion), dann nimmt die Ausprägung von AV1 (Akzeptanz) ab.

2. Stichprobengröße

Wie aus der statistischen Formulierung von h6 ersichtlich, wird für h6 mit Korrelationen getestet. Die Korrelation betrifft zwei abhängige Variablen. Daher kann auf den Manipulationscheck verzichtet und mit $N = 598$ Probanden getestet werden.

3. Hypothesenprüfung (Korrelationen)

In Szenario B, C und D wird die Nullhypothese mit $p < 0.1\%$ falsifiziert (vgl. Tab. 6-7). In diesen Szenarien besteht eine signifikante, gleichsinnige Korrelation zwischen der emotionalen Reaktion und dem Akzeptanzniveau (r (B) = .39; r (C) = .020; r (D) = .58, $df_{B,C,D} \geq 100$). In Szenario A wird die Nullhypothese nicht falsifiziert. Szenario A weist mit einem r -Wert von .04 eine marginale Korrelation zwischen der emotionalen Reaktion und der Akzeptanz auf. Über alle Szenarien summiert ist die Korrelation auf dem $p < 1\%$ Niveau signifikant, mit einer Korrelation von $r = .12$ und $df \geq 100$.

Tabelle 6-7: Zusammenhang zwischen emotionaler Reaktion und Akzeptanz

h6: AV4 → AV1		Korrelationen (<i>n</i> = 598)		
Szenario	<i>r</i>	Sig. (2-seitig)	<i>d</i>	
A. Übergepäck	.04	.30	0.09	
B. Lounge / VIP	.39	.00	0.85	
C. Rückreisedatum	.20	.00	0.40	
D. Meilen-Upgrade	.58	.00	1.41	
Total (Szenario unabhängig)	.12	.01	0.24	

Quelle: Eigene Berechnung mit SPSS und G*Power

4. Effektgrösse und Teststärke

Mit $d = 0.24$ ist die Effektgrösse über alle Szenarien hinweg betrachtet (A + B + C + D), schwach ausgeprägt. Bei Einzelbetrachtung der Szenarien besteht im Gegensatz zum Gesamteffekt ein starker Effekt in B ($d = 0.85$) und ein sehr starker Effekt in Szenario D ($d = 1.41$). Der schwache Gesamteffekt wird insbesondere von Szenario A ($d = 0.09$) beeinflusst (vgl. Tab. 6-7).

Die Teststärke ist in Szenario B und D stark ausgeprägt mit einem Wert von $(1 - \beta) = 1$. In Szenario A und C ist die Teststärke tiefer, mit den Werten $(1 - \beta) = .70$ und $(1 - \beta) = .87$.

5. Interpretation und Übersicht (h6)

Insgesamt besteht zwischen Emotion und Akzeptanz eine signifikante, gleichsinnige Korrelation. Zwischen den Variablen *Emotion* und *Akzeptanz* wird ein starker Effekt in Szenario B (Lounge-Zugang) und D (Meilen-Upgrade) nachgewiesen. Dieser Effekt wird mit einer tiefen Irrtumswahrscheinlichkeit und grosser Teststärke nachgewiesen.

Die schwache Korrelation im Übergepäck-Szenario (A) lässt vermuten, dass sich Passagiere bereits an dieses Ereignis gewöhnt haben. Eine Übergepäckgebühr löst wenig bis keine emotionale Reaktion aus. Das Akzeptanzniveau bleibt relativ hoch. In Szenario C (Rückreisedatum) handelt es sich ebenfalls um einen Sachverhalt, den Passagiere kennen oder bereits angetroffen haben. Daher wird möglicherweise in C eine signifikante, aber im Vergleich zu B und D tiefere Korrelation ($r = .20$) zwischen Emotion und Akzeptanz gemessen.

6.3.7 Akzeptanzniveau und negative Verhaltensabsichten (h7)

1. Formulierung Hypothese 7

h7 modelltheoretisch formuliert: Wenn das Akzeptanzniveau (AV1) tief ist (Kundensteuerung mit Richtlinien), dann steigt die Wahrscheinlichkeit von negativen Verhaltensabsichten (AV2_1 Beschwerde, AV2_2 negative WOM, AV2_3 Illoyalität).

2. Stichprobengröße

Für die Prüfung negativer Verhaltensabsichten wird die Faktorstufe UV1 „Richtlinie“ verwendet, da aus inhaltlichen Gründen nur unter dieser Faktorstufe negative Verhaltensabsichten erwartet werden. Diese Aussage wird im zweiten Experiment empirisch überprüft (vgl. Kap. 6.5.1). Der Hypothesentest zur Manipulation hat in h1 für UV1 bestätigt, dass alle Probanden unter der Faktorstufe „Richtlinie“ verwendet werden können (vgl. Kap. 6.3.1). Dies ergibt eine Stichprobengröße von $n = 275$ für h7.

3. Hypothesenprüfung (Korrelationen)

Zur Hypothesenprüfung von h7 wird mit den Anforderungen an das Signifikanzniveau und an die Korrelationsgröße sowie -richtung ein strenges Kriterium verwendet (vgl. Tab. A-2):

H_1 : $r \leq -.11$ (spezifisch, gegensinnig für AV2_1 und AV2_2). Für AV2_3 gilt: $r \geq .11$ (spezifisch, gleichsinnig). $df = 273$ (d.h. für r approximativer Wert bei $df = 300$ verwendet).

In *Szenario A* wird die Alternativhypothese für „Beschwerde“ angenommen ($p < 0.1\%$). Der geforderte Korrelationswert und die Wirkungsrichtung ($r(A) = -.27$, $n = 275$, $p = .00$) werden erfüllt (vgl. Tab. 6-8). Bei „neg. WOM“ wird die Nullhypothese nicht falsifiziert, da $p = .24$. Für die „Illoyalität“ besteht eine auf dem $p < 0.1\%$ Niveau signifikante Beziehung, mit $r = .18$. Die Hypothese und die Wirkungsrichtung stimmen, weshalb die Alternativhypothese angenommen wird.

Tabelle 6-8: Zusammenhang zwischen Akzeptanz und negativen Verhaltensabsichten

h7: AV1 → AV2_1-3		Korrelationen ($n = 275$)		
Szenario	Verhalten	r	Sig. (2-seitig)	d
A. Übergepäck	Beschwerde	-.27	.00	0.55
	Negative WOM	-.07	.24	-
	Illoyalität	.18	.00	0.37
B. Lounge / VIP	Beschwerde	.03	.62	-
	Negative WOM	.11	.08	-
	Illoyalität	.26	.00	0.54
C. Rückreisedatum	Beschwerde	-.22	.00	0.45
	Negative WOM	-.20	.00	0.41
	Illoyalität	.33	.00	0.70
D. Meilen-Upgrade	Beschwerde	-.51	.00	1.20
	Negative WOM	-.37	.00	0.80
	Illoyalität	.49	.00	1.12

Quelle: Eigene Berechnung mit SPSS und G*Power

In *Szenario B* erreicht die Beschwerde das geforderte Signifikanzniveau nicht ($p = .62$). Bei „neg. WOM“ wird die Nullhypothese nicht falsifiziert, da die Wirkungsrichtung der Korrelation nicht erfüllt ist ($r = .11$). Dagegen wird die Alternativhypothese für „Illoyalität“ angenommen ($r(B) = .26$,

$n = 275$, $p = .00$). In *Szenario C* werden alle Korrelationsrichtungen (vgl. oben) und die Anforderungen an die Signifikanz erfüllt. In *Szenario D* wird die Alternativhypothese für alle negativen Verhaltensabsichten angenommen, auf dem hohen Signifikanzniveau von $p < 0.1\%$. Am stärksten ausgeprägt sind die Korrelationen unter Szenario D, „Beschwerde“ mit einem r -Wert von $-.51$, resp. $r = .49$ für „Illoyalität“.

Aussagekräftig ist die bei allen Szenarien bestehende hohe Korrelation zwischen „Beschwerde“ und „neg. WOM“ mit einem r -Wert von $.41$ (A) bis $.54$ (B) ($n = 275$, $p = .00$ bis $.04$). Auch ausgeprägt ist das Verhältnis zwischen „Beschwerde“ und „Illoyalität“, wobei die Korrelation gegensinnig und etwas schwächer ausfällt, mit r -Werten von $-.30$ (B) bis $-.44$ (D) ($n = 275$, $p = .00$ bis $.03$). Die geforderte Anzahl Freiheitsgrade (df) wird in allen getesteten Hypothesen erfüllt.

4. Effektgrösse und Teststärke

Ein r -Wert von $.26$ in Szenario B („Illoyalität“) bedeutet ein Differenzmass von $d = 0.54$ (mittlerer Effekt). Ebenfalls ein mittlerer Effekt ist in Szenario A die „Beschwerde“ mit einem Differenzmass von $d = 0.55$. Starke Effekte werden in Szenario C für das illoyale Verhalten ($d = 0.70$) und für alle negativen Verhaltensabsichten unter Szenario D nachgewiesen ($d \geq 0.8$).

Die Teststärke der angenommenen Hypothesen unter A („Beschwerde“) liegt bei $(1 - \beta) = .56$, in B („Beschwerde“) bei $(1 - \beta) = .35$ und in C („illoyales Verhalten“) sowie D (alle Verhaltensabsichten) bei $(1 - \beta) > .90$.

5. Interpretation und Übersicht (h7)

In Szenario A (Übergepäck) wird die Hypothese für die „Beschwerde“ und „illoyales Handeln“ bestätigt. Im Gegensatz dazu wird die Alternativhypothese für „neg. WOM“ falsifiziert. Wie unter h_6 ist Übergepäck ein den Passagieren bekanntes Phänomen. Die negativen Verhaltensabsichten sind daher nicht stark ausgeprägt.

In Szenario B (Lounge-Zugang) wird die Alternativhypothese bezüglich „Illoyalität“ bestätigt. Passagiere interpretieren einen fehlenden Lounge-Zugang als ein Ereignis, für das sich beschweren oder „neg. WOM“ nicht lohnt. Sie wählen die Verhaltensabsicht, welche am wenigsten „Einsatz“ verlangt.

In Szenario C (Rückreisedatum) und D (Meilen-Upgrade) kann die Hypothese bezüglich aller negativen Verhaltensabsichten bestätigt werden (mit hoher Signifikanz und Teststärke). Je tiefer das Akzeptanzniveau bei einem negativen Ereignis, umso höher ist die Wahrscheinlichkeit von „Beschwerde“, „negative WOM“ und „Illoyalität“. Dabei sind die Absicht, sich zu beschweren, und illoyales Handeln am deutlichsten ausgeprägt.

Probanden, die sich beschweren, werden sich, über alle Szenarien hinweg betrachtet, illoyal verhalten und Freunden vom negativen Ereignis erzählen.

6.3.8 Demografische und abhängige Variablen (h_d)

Der Einfluss demografischer Variablen wird unabhängig von den experimentellen Bedingungen erhoben und als Demografie-Nebenhypothese (h_d) geprüft. Im Kontext der forschungsleitenden Fragestellung sind vor allem die Einflüsse von Variablen des Flugverhaltens auf die abhängigen Variablen *Akzeptanz* und *negative Verhaltensabsichten* von Interesse. Untersucht wird der Einfluss der Klassenzugehörigkeit (Produkt: *Economy*, *Business* und *First Class*), Anzahl Flüge im Jahr 2006 und Status des Passagiers im Vielfliegerprogramm. Zur statistischen Überprüfung werden die in Kapitel 4.2 inhaltlich-theoretisch formulierten Hypothesen h_{d1} und h_{d2} mit je zwei Sub-Hypothesen (a, b) getestet.

Klassenzugehörigkeit und Akzeptanz (h_{d1a})

Mit einem *t*-Test für Mittelwert-Gruppenunterschiede wird die Hypothese getestet, dass eine tiefere Klassenzugehörigkeit zu einer tieferen Akzeptanz des Kundensteuerungskonzeptes führt und die Wahrscheinlichkeit negativer Verhaltensabsichten steigt. Der Datensatz der *First Class* ist zu tief, um einen Vergleich durchzuführen ($n = 2$). Daher wird zwischen *Economy* ($n = 549$) und *Business Class* ($n = 41$) verglichen. Die interessierenden abhängigen Variablen sind das Akzeptanzniveau und die negativen Verhaltensabsichten.

Tabelle 6-9: Auswirkung der Klassenzugehörigkeit auf Akzeptanz

h_{d1a} : Klasse \rightarrow AV1		Klasse		t-Test (2-seitig)			Effekt
		<i>Economy</i> $n = 549$	<i>Business</i> $n = 41$	<i>df</i>	<i>t</i>	<i>Sig.</i>	<i>d</i>
A	Mittelwert	2.82	2.92	578	-0.30	.76	0.05
	Stand.abw.	1.90	1.85				
B	Mittelwert	2.11	3.08	584	-3.50	.00	0.52
	Stand.abw.	1.63	2.10				
C	Mittelwert	2.77	2.55	578	0.71	.48	0.12
	Stand.abw.	1.78	1.90				
D	Mittelwert	3.21	3.97	579	-2.42	.02	0.41
	Stand.abw.	1.88	1.82				
Σ	Mittelwert	10.88	12.38	564	-2.26	.02	0.36
	Stand.abw.	3.84	4.53				

Quelle: Eigene Berechnung mit SPSS und G*Power

In der Gruppenstatistik fallen Szenario B (Lounge-Zugang) und D (Meilen-Upgrade) mit grossen Mittelwertsunterschieden auf (vgl. Tab. 6-9 für die relevanten *t*-Test Werte). Ein *Economy-Class*-Passagier in Szenario B bewertet die Ereignisse auf einer Akzeptanzskala von 1 (tiefe Akzeptanz) bis 6 (hohe Akzeptanz) mit einem durchschnittlichen Wert von $\bar{x}_{Eco} = 2.11$ ($s = 1.63$). Der *Business-Class*-Passagier hingegen akzeptiert den Sachverhalt mit $\bar{x}_{Bus} = 3.08$ ($s = 2.10$) stärker. Der Unterschied ist signifikant ($p(B) = .00$; $p(D) = .02$). In Szenario A (Übergepäck) und C (Rückreisedatum)

sind die Unterschiede nicht signifikant, was der t -Test für Mittelwertsvergleiche bestätigt ($p(A) = .76$; $p(D) = .48$).

In der Totalbetrachtung über alle Szenarien unterscheiden sich *Economy*- von *Business-Class*-Passagieren signifikant bezüglich des Akzeptanzniveaus ($p < 5\%$ Niveau). Die Nullhypothese wird falsifiziert: *Economy-Class*-Passagiere weisen ein tieferes Akzeptanzniveau als *Business-Class*-Passagiere auf mit einer Teststärke von $(1 - \beta) = .44$.

Klassenzugehörigkeit und Verhaltensabsichten (h_{d1b})

Bei den negativen Verhaltensabsichten bestehen über alle Szenarien betrachtet nur schwache Unterschiede zwischen *Economy*- und *Business-Class*-Passagieren, trotz unterschiedlichen Akzeptanzniveaus. Die Nullhypothese, dass sich *Economy*- und *Business-Class*-Passagiere nicht unterscheiden, wird in Szenario A, C und D nicht falsifiziert.

Tabelle 6-10: Auswirkung der Klassenzugehörigkeit auf Verhaltensabsichten

h_{d1b} : Klasse \rightarrow AV2_1 -3			Klasse		t-Test (2-seitig)			Effekt
			<i>Economy</i> $n = 542$	<i>Business</i> $n = 38$	<i>df</i>	<i>t</i>	Sig.	<i>d</i>
B.	Beschwerde	Mittelwert	2.34	2.92	580	-2.08	.04	0.32
		Stand.abw.	1.63	2.01				
B.	Illoyalität	Mittelwert	2.83	3.31	572	-2.06	.04	0.30
		Stand.abw.	1.37	1.81				

Quelle: Eigene Berechnung mit SPSS mit G*Power

Wird der t -Test in Szenario B (Lounge-Zugang) für alle negativen Verhaltensabsichten durchgeführt, wird ein Mittelwertsunterschied zwischen *Economy* und *Business Class* ersichtlich (vgl. Tab. 6-10). *Economy*- und *Business-Class*-Passagiere haben unterschiedlich ausgeprägtes Beschwerdeverhalten ($p = .04$, $d = 0.32$, $(1 - \beta) = .43$). Ebenfalls signifikant unterscheidet sich in Szenario B die Verhaltensabsicht „Illoyalität“ zwischen den Klassen ($p = .04$, $d = 0.30$, $(1 - \beta) = 0.40$). *Economy-Class*-Passagiere verhalten sich illoyaler als *Business-Class*-Passagiere. Trotz eines durchschnittlich tieferen Akzeptanzniveaus beschwerten sich in Szenario B mehr *Business*- als *Economy-Class*-Passagiere. In Tabelle 6-10 sind aus Platzgründen nur die signifikanten Resultate aufgeführt.

Flugerfahrung und Akzeptanz sowie Verhaltensabsichten (h_{d2a-b})

Wird die *Anzahl Geschäftsflüge* auf eine Korrelation mit dem Summenwert der Akzeptanz untersucht, ergibt dies einen r -Wert von .07, der nicht signifikant ist ($n = 245$, $p = .70$). Bezüglich Verhaltensabsichten hat die „Flugerfahrung“ bei geschäftlich Reisenden lediglich in Szenario B (Lounge-Zugang) eine signifikante Korrelation von $r = .14$ ($n = 245$, $p = .04$). Je weniger Flugerfahrung ein geschäftlich Reisender hat, umso „illoyaler“ verhält er sich (vgl. Tab. 6-11).

Die Korrelation zwischen der *Anzahl Privatflüge* und dem Akzeptanzniveau, über alle Szenarien hinweg betrachtet, ist mit einem r -Wert von .15 gleichsinnig, mittelstark ausgeprägt ($n = 478$, $p = .00$). Am deutlichsten tritt diese Korrelation in Szenario C auf: Je erfahrener ein Privatreisender ist, umso eher akzeptiert er, dass er keinen Lounge-Zugang hat ($r = .12$, $n = 478$, $p = .01$). Die Anzahl Privatflüge korreliert nicht mit den negativen Verhaltensabsichten. In Tabelle 6-11 sind aus Platzgründen nur die signifikanten Resultate aufgeführt. „Summe“ bedeutet, dass über alle Szenarien gerechnet wurde (A + B + C + D), „Einzel“ bezeichnet das signifikanteste Szenario.

Tabelle 6-11: Auswirkung der Flugerfahrung auf Akzeptanz und Verhaltensabsichten

		Akzeptanz			Neg. Verhaltensabsichten		
		Szenario	r	Sig. (2-seitig)	Szenario	r	Sig. (2-seitig)
h_{d2a} : Flugerfahrung → AV1							
h_{d2b} : Flugerfahrung → AV2_1-3							
Anzahl Privatflüge	Summe	Alle	.15	.00	Alle	-	n.s.
	Einzel	C	.12	.01	-	-	-
Anzahl Geschäftsflüge	Summe	Alle	-	n.s.	Alle	-	n.s.
	Einzel	-	-	-	B: Illoyalität	.14	.04
Mitgliedschaft	Summe	Alle	-	n.s.	Alle	-	n.s.
	<i>Gold Member</i>	-	-	-	A: Beschwerde	-	.00

Quelle: Eigene Berechnung mit SPSS

Neben der Anzahl Flüge wird die Flugerfahrung durch die Teilnahme am Vielfliegerprogramm mitbestimmt. In Bezug auf das Akzeptanzniveau und die negativen Verhaltensabsichten gibt es keinen signifikanten Zusammenhang zwischen Mitgliedern eines Vielfliegerprogrammes und Nichtmitgliedern. Der Status im Vielfliegerprogramm hat lediglich einen Einfluss in Szenario A (Übergepäck): Die *Gold Members* (zweithöchste Stufe) beschwerten sich eher als die anderen Mitglieder im Vielfliegerprogramm ($F = 4.39$, $df = 578$, $n = 579$, $p = .00$).

6.4 Prüfung Forschungsmodell Experiment I

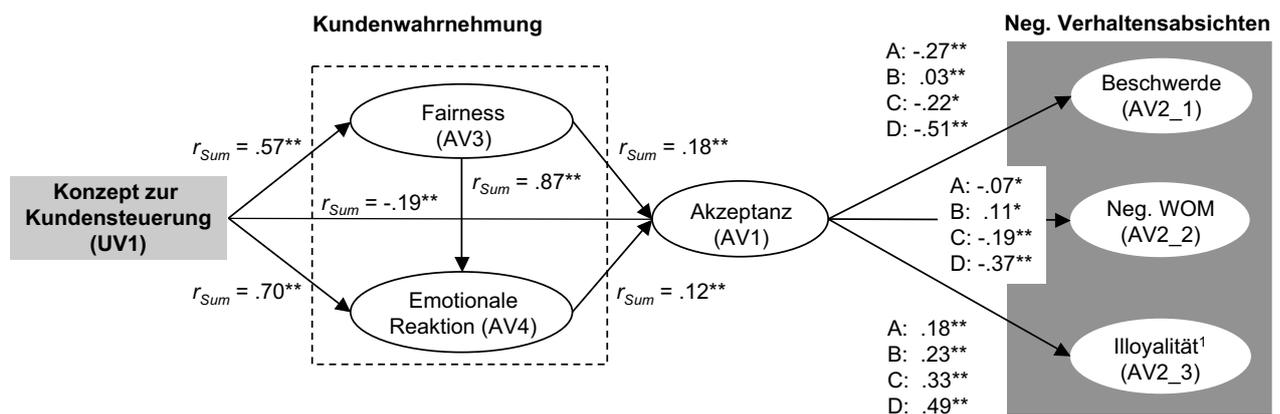
Mit Kapitel 6.4.1 werden im zentralen Forschungsmodell die Korrelationen zwischen den Variablen abgebildet. Untersucht werden Korrelationen, die je nach Kundensteuerung auf die einzelnen Szenarien aufgeschlüsselt werden. Die Prüfung der Interaktionen im Modell wird in Kapitel 6.4.2 durchgeführt. Die Anpassung des zentralen Forschungsmodells (Kap. 6.4.3) im Hinblick auf die Online-Untersuchung beruht auf Erkenntnissen vom letzten Kapitel. Um die Anpassungen empirisch testen zu können, muss der Online-Fragebogen angepasst werden - die durchgeführten Modifikationen werden in Kapitel 6.4.4 vorgestellt.

Aufgrund des insgesamt schwachen Zusammenhangs zwischen den Probandendaten und dem Antwortverhalten wird der Zusammenhang zwischen endogenen demografischen und den abhängigen Variablen im zweiten Experiment nicht weiter untersucht.

6.4.1 Korrelationen im Kundensteuerungskonzept

Um das Gesamtmodell mit Korrelationen abzubilden, wird von der Hauptvariablen *Kundensteuerung* ausgegangen. Die in Abbildung 6-4 aufgeführten Korrelationen (Pfeile mit Wirkungsrichtung) beziehen sich auf die Korrelationssumme über alle Szenarien A bis D hinweg, ausser zwischen der Variablen *Akzeptanz* und *negative Verhaltensabsichten* ($n = 598$). Die geforderte Anzahl Freiheitsgrade ist in allen Szenarien erfüllt.

Abbildung 6-4: Korrelationen im Modell in Bezug auf Kundensteuerung



Legende: * = Korrelation auf $p < .05$ Niveau, ** = Korrelation auf $p < .01$ Niveau, ¹ = Illoyalität ist gegensinnig skaliert;

Quelle: Eigene Berechnung mit SPSS

Die Korrelationswerte zwischen *Kundensteuerung-Fairness* und *Kundensteuerung-Emotion* sowie *Fairness-Akzeptanz* und *Emotion-Akzeptanz* zeigen, dass die Fairness-Beurteilung das Akzeptanzniveau eines bestimmten Sachverhalts ähnlich gut vorhersagen kann wie das Konstrukt *Emotion*.

Widersprüchlich erscheint in diesem Zusammenhang die gegenseitige Korrelation zwischen Kundensteuerung (UV1) und Akzeptanz (AV1). Werden Probanden gefragt, wie fair sie einen Sachverhalt einstufen, unterscheidet sich das Antwortverhalten vom konkreten Verhalten („Bezahlen Sie die Gebühr?“). Über alle Szenarien betrachtet, bezahlen Probanden die Gebühr am Check-In (Richtlinie), obwohl sie den Sachverhalt als unfair einstufen und obwohl mit der Option am Y-Air Büro kein „Zwang“ besteht. Wird der Korrelationskoeffizient ($r_{Sum} = -.19$) pro Szenario für die direkte Beziehung *Kundensteuerung-Akzeptanz* aufgeschlüsselt, wird im Modell ein neuer Zusammenhang erkennbar: In den Szenarien A und C ($r(A) = -.44$, $r(C) = -.30$) besteht eine gegenseitige (negative) und in den Szenarien B und D eine gleichsinnige (positive) Beziehung zwischen Kundensteuerung und Akzeptanz (vgl. Tab. 6-12).

Tabelle 6-12: Direkte Auswirkung der Kundensteuerung auf Akzeptanz der Gebühr

UV1 → AV1		Kundensteuerung		Effekt	P.b. Korrelation
Szenario		Richtlinie (tief) <i>n</i> = 275	Anreiz (hoch) <i>n</i> = 323	<i>d</i>	<i>r</i> Total: -.19
A. Übergepäck	Mittelwert	3.72	2.06	1	-.44
	Stand.abw.	1.91	1.50		
B. Lounge / VIP	Mittelwert	1.97	2.35	0.23	.11
	Stand.abw.	1.56	1.76		
C. Rückreise	Mittelwert	3.33	2.26	0.62	-.30
	Stand.abw.	1.84	1.59		
D. Meilen-Upgrade	Mittelwert	2.76	3.67	0.50	.24
	Stand.abw.	1.69	1.92		

Quelle: Eigene Berechnung mit SPSS und G*Power

Die Steuerung über Online-Anreize ist demnach für den Lounge-Zugang sowie Meilen-Upgrade geeignet (vgl. h1). Die gegenseitige Beziehung in Szenario A und C bedeutet, dass Probanden am Check-In eine Übergepäckgebühr (A) oder ein flexibles Ticket (C) mit hoher Wahrscheinlichkeit bezahlen. Bei einer Online-Buchung hingegen verzichten die Probanden auf diese Anreiz-Optionen. Dies widerspricht der Fairness-Beurteilung, welche eine gleichsinnige Korrelation zwischen *Kundensteuerung-Fairness-Beurteilung-Akzeptanz* postuliert. Probanden stufen demnach eine Übergepäckgebühr und eine Gebühr für ein flexibles Ticket am Check-In als unfair ein, aber sie bezahlen diese Zusatzkosten. Zwischen dem Verhalten (Akzeptanz) und der Einstellung (Fairness-Beurteilung) wird eine Diskrepanz nachgewiesen.

Werden die aufsummierten Korrelationskoeffizienten auf einzelne Szenarien und experimentelle Bedingungen aufgeschlüsselt, wird deutlich, dass die Korrelationswerte teilweise stark schwanken (vgl. Tab. 6-13). Zur theoretischen Modellprüfung eignen sich die summierten Werte. Für praktische Implikationen müssen jedoch die Werte, aufgeschlüsselt nach Szenario, berücksichtigt werden. Unter Einbezug der Korrelationsrichtung (gleichsinnig, gegensinnig) und Effektstärke wird ersichtlich,

dass die Szenarien B (Lounge-Zugang) und D (Meilen-Upgrade) die deutlichsten Ausprägungen im Forschungsmodell aufweisen. Für die Online-Untersuchung eignen sich daher die Szenarien B und D am besten.

Tabelle 6-13: Korrelation, Signifikanz und Effektstärke in Szenarien A bis D

	Kundensteuerung	Kommunikation	Ticketpreis	Image
	Hypothese 1	Hypothese 2	Hypothese 3	Hypothese 4
A. Übergepäck	$r = .18$ $p = .00$ $d = 0.37$ (mittel)	$r = .22$ $p = .00$ $d = 0.46$ (mittel)	$r = .05$ $p = .41$ -	$r = .16$ $p = .01$ $d = 0.32$ (mittel)
B. Lounge / VIP	$r = .29$ $p = .00$ $d = 0.60$ (mittel)	$r = .28$ $p = .00$ $d = 0.61$ (mittel)	$r = .11$ $p = .06$ -	$r = .35$ $p = .00$ $d = 0.76$ (stark)
C. Rückreisedatum	$r = .38$ $p = .00$ $d = 0.82$ (stark)	$r = .19$ $p = .02$ $d = 0.40$ (mittel)	$r = .08$ $p = .14$ -	$r = .27$ $p = .00$ $d = 0.57$ (mittel)
D. Meilen-Upgrade	$r = .70$ $p = .00$ $d = 1.94$ (stark)	$r = .24$ $p = .00$ $d = 0.50$ (mittel)	$r = -.01$ $p = .86$ -	$r = .34$ $p = .00$ $d = 0.73$ (stark)

□ = Hypothese bestätigt ▨ = Hypothese abgelehnt

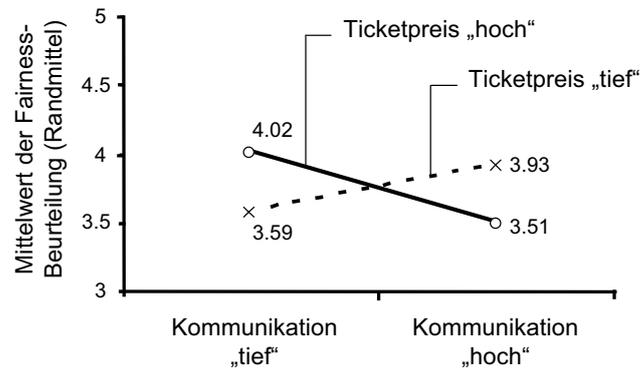
Quelle: Eigene Berechnung mit SPSS und G*Power

6.4.2 Interaktions-Hypothesen (h_i)

Das A-priori-Forschungsmodell für das Experiment I postuliert, dass die unabhängigen Variablen *Ticketpreis* (UV3) und *ex ante Image* (UV4) einen moderierenden Effekt auf die Wirkbeziehung von UV1 (Kundensteuerung) und UV2 (Kommunikation) auf die Fairness-Beurteilung (AV3) haben. Mit der univariaten, zweifaktoriellen Varianzanalyse lässt sich dieser Interaktionseffekt zweier UVn mit dem *F*-Test prüfen (Hirsig, 2002).

Zur Analyse der Interaktionseffekte werden nur die Stichproben verwendet, welche den Manipulationscheck bezüglich Ticketpreis (Image) bestanden haben. Die Interaktions-Hypothesen werden mit der univariaten Varianzanalyse für jedes Szenario und jede Interaktions-Hypothese einzeln geprüft. Detaillierte Forschungsergebnisse sind im Anhang aufgeführt (A.5.3).

Obwohl der Faktor *Ticketpreis* nicht signifikant mit der kognitiven Fairness-Beurteilung korreliert (vgl. h_3), wird in Szenario D (Meilen-Upgrade) ein Interaktionseffekt mit den Faktoren Kundensteuerung ($F = 5.58$, $df = 326$, $n_{tief} = 183$, $n_{hoch} = 147$, $p = .02$) sowie Kommunikation ($F = 4.48$, $df = 326$, $n_{tief} = 183$, $n_{hoch} = 147$, $p = .04$) nachgewiesen. Verglichen werden die Mittelwerte beider UV und Faktorstufen „tief“ sowie „hoch“ (vgl. Abb. 6-5). In den Szenarien A, B und C wird die Alternativhypothese zur Interaktion zwischen Ticketpreis und Kundensteuerung / Kommunikation als nicht-signifikant abgelehnt (vgl. A.5.3). Die disordinale Interaktion zwischen Ticketpreis und Kommunikation wird in Abbildung 6-5 grafisch dargestellt.

Abbildung 6-5: Interaktion zwischen Ticketpreis und Kommunikation (Szenario D)

Quelle: Eigene Darstellung

Der Faktor *ex ante Image* interagiert weder mit der Kundensteuerung noch mit Kommunikation. In keinem Szenario konnte diese Interaktion nachgewiesen werden. Die Fairness-Beurteilung als AV wird unabhängig vom *ex ante Image* und von der Kundensteuerung und deren Kommunikation beeinflusst. Im Modell wird die Variable *ex ante Image* nicht als Moderator bestätigt; sie hat eine unabhängige Auswirkung auf die Fairness-Beurteilung.

6.4.3 Adaption des Modells für Experiment II

Das Forschungsmodell wird mit dem ersten Experiment bestätigt. Vor allem in Szenario B (Lounge-Zugang) und D (Meilen-Upgrade) wird das Modell über die experimentellen Bedingungen *Kundensteuerung*, *Kommunikation* und *ex ante Image* deutlich bestätigt. Die experimentelle Bedingung *Ticketpreis* hat keine direkte Auswirkung auf die Fairness-Beurteilung. Trotzdem hat der Ticketpreis einen leicht gegensinnig, moderierenden Effekt auf die Kundensteuerung und die Kommunikation. *Ex ante Image* hingegen ist keine Moderator-Variable. Sie beeinflusst die Fairness-Beurteilung unabhängig von anderen experimentellen Bedingungen.

Die Variable *Ticketpreis* wird im zweiten Experimente nicht weiter untersucht, weil sie implizit durch das Geschäftsmodell in der Versuchsanordnung definiert ist: Ein *Low Cost Carrier* zeichnet sich im Vergleich zum *Network Carrier* durch einen tieferen Preis aus.

Neben den bereits formulierten Hypothesen für das zweite Experiment (h1, h2, h7-h9) berücksichtigt das zweite (Online-) Experiment die statistischen Auswertungen des ersten Experiments. Dieses sequentielle Vorgehen ermöglicht die Anpassung des Modells und Einbezug neuer Erkenntnisse in das A-priori-Forschungsmodell für die Online-Umfrage.

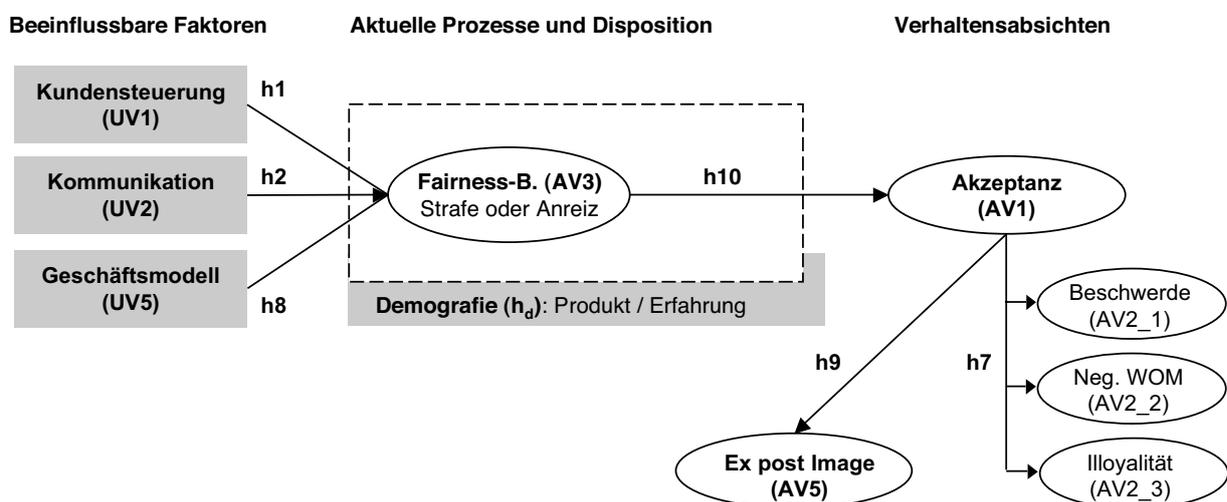
Die Untersuchung der Korrelationen aus der Flughafen-Umfrage zeigt, dass zwischen der Einstellung und dem Verhalten der Probanden eine Diskrepanz besteht. Zudem verdeutlichen die Korrelationskoeffizienten, dass die Hypothesen bezüglich *emotionaler Reaktion* für das zweite Experiment

angepasst werden müssen (vgl. h5 und h6). Zwischen der kognitiven Fairness-Beurteilung und der Akzeptanz wird eine direkte Verbindung postuliert. Die „Synthese“ aus h5 und h6 wird neu als Hypothese 10 (h10) in der Online-Umfrage untersucht und im Forschungsmodell einbezogen. Die inhaltliche Formulierung lautet:

h10: Je höher die Fairness-Beurteilung, umso höher ist die Akzeptanz des Kundensteuerungskonzeptes.

Das bezüglich h10 angepasste 2x2x4 Forschungsmodell des Online-Experiments ist in Abbildung 6-6 dargestellt. Es basiert auf dem zentralen, A-priori-Forschungsmodell und beinhaltet die relevanten Wirkungsbeziehungen für das web-gestützte Experiment (vgl. Kap. 4.2).

Abbildung 6-6: A-priori-Forschungsmodell Experiment II



Quelle: Eigene Darstellung

6.4.4 Anpassung des Fragebogens für Online-Experiment

Länge und Komplexität des Online-Fragebogens konnten aufgrund der Erkenntnisse der Flughafen-Umfrage sowie anhand der neuen Fragestellungen reduziert werden, weil:

- keine detaillierte Einzelbewertung der negativen Verhaltensabsichten nötig war, da „illoyales Verhalten“ als ausgeprägteste Verhaltensreaktion identifiziert wurde. Pro Verhaltensabsicht wurde nicht mehr eine eigene Skala verwendet. Es wurde nach der „wahrscheinlichsten“ Reaktion gefragt, wobei Mehrfachantworten möglich waren;
- nur drei Manipulationen der UVn getestet wurden. Die Moderatorvariable *Ticketpreis* wurde nicht weiter untersucht. Die UV *ex ante Image* wurde neu als abhängige Variable (*ex post*) untersucht;

- die Messung des Konstrukts *emotionale Reaktion* erübrigte sich (Hypothese 10, vgl. oben);
- anstatt vier nur noch zwei Szenarien untersucht wurden. Aufgrund der deutlichen Ergebnisse im ersten Experiment werden Szenario B (Lounge-Zugang) und D (Meilen-Upgrade) gewählt;
- die Skalierung des UV1-Manipulationschecks wurde mit einer 3er Likert-Skala vereinfacht;
- bezüglich Flugverhalten eine weniger detaillierte Abfrage nötig ist. Da sich Probanden während der Online-Befragung nicht im Airline-Serviceprozess befinden, sind detaillierte Fragen zum letzten Flug schwierig zu beantworten. Im Gegensatz zum schriftlichen Fragebogen wurden die Fragen zum Flugverhalten auf das Jahr 2006 bezogen.

6.5 Hypothesentest Experiment II

Im Online-Experiment (Experiment II) werden mit h1, h2 und h7 dieselben statistischen Hypothesen wie im ersten Experiment getestet. Da es sich um eine Replikation handelt, wird im Kapitel 6.5.1 lediglich die Interpretation der Ergebnisse aufgeführt (detaillierte Ergebnisse sind in A.5.4). Hypothese 8 betrifft das Geschäftsmodell und ist in Kapitel 6.5.2 vorgestellt. Mit Hypothese 9 wird der Einfluss des Akzeptanzniveaus auf die ex post Image-Beurteilung untersucht (Kap. 6.5.3). In Kapitel 6.5.4 wird Hypothese 10 überprüft, die aufgrund von Erkenntnissen des ersten Experiments erstellt wurde.

Bei der Auswertung der Online-Umfrage steht die detaillierte Analyse aufgeschlüsselt nach Szenario B (Lounge-Zugang) und D (Meilen-Upgrade) im Vordergrund. Korrelationen im Forschungsmodell, demografische Einflüsse und Interaktionen der UVn werden nicht untersucht. Die Auswertung der Flughafen-Umfrage hat hierzu eindeutige und umfangreiche Daten geliefert. Eine Übersicht zu den Resultaten der UV-Manipulationen im Online-Experiment sind in Tabelle 6-14 dargestellt.

Tabelle 6-14: Korrelation, Signifikanz und Effektstärke in Szenarien B und D

	Kundensteuerung	Kommunikation	Geschäftsmodell
	Hypothese 1	Hypothese 2	Hypothese 8
B. Lounge-Zugang	$t = -18.17$ $p = .00$ $d = 1.96$ (stark)	$t = -7.96$ $p = .00$ $d = 0.84$ (stark)	$F = 1.68$ $p = .17$ -
D. Meilen-Upgrade	$t = -34.72$ $p = .00$ $d = 2.93$ (stark)	$t = -8.31$ $p = .00$ $d = 0.60$ (mittel)	$F = 0.64$ $p = .59$ -

 = Hypothese bestätigt  = Hypothese abgelehnt

Quelle: Eigene Berechnung mit SPSS und G*Power

6.5.1 Replizierte Hypothesen (h1, h2 und h7)

Kundensteuerung und Fairness-Beurteilung (h1)

Aufgrund des *t*-Tests zum Vergleich von Mittelwerten unabhängiger Stichproben wird in Szenario B (Lounge-Zugang) und D (Meilen-Upgrade) die Hypothese bestätigt. Der Unterschied zwischen der Fairness-Beurteilung unter der Faktorstufe „Richtlinie“ unterscheidet sich für denselben Sachverhalt signifikant von der Faktorstufe „Anreiz“. Es spielt bei der Fairness-Beurteilung von Passagieren eine Rolle, ob das Konzept zur Kundensteuerung als Richtlinie oder als Anreiz gestaltet wird. Die hohe Signifikanz und Teststärke verdeutlichen, dass dieser Zusammenhang nicht zufällig entstanden ist. Die Fehler erster und zweiter Art sind minimal.

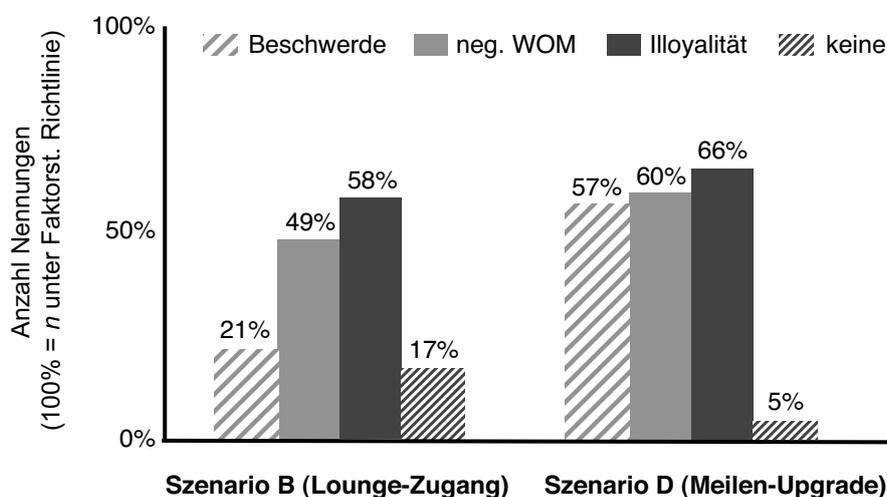
Kommunikation und Fairness-Beurteilung (h2)

Die Alternativhypothese von h2 wird für beide Szenarien B (Lounge-Zugang) und D (Meilen-Upgrade) mit hoher Signifikanz und Teststärke angenommen: Je deutlicher und klarer die Anreiz-Optionen sowie Konditionen bei der Online-Buchung kommuniziert werden, umso „fairer“ wird eine Tarifoption oder Bedingung eingestuft. Der Effekt ist in beiden Szenarien deutlich ausgeprägt: Das Design der Webpage für die Online-Buchung beeinflusst die Kundenwahrnehmung.

Akzeptanzniveau und negative Verhaltensabsichten (h7)

Ein tiefes Akzeptanzniveau führt zu negativen Verhaltensabsichten. Die wahrscheinlichste Reaktion, wenn die Kundensteuerung über Richtlinien erfolgt, ist in dieser Untersuchung „Illoyalität“, gefolgt von „negativer WOM“. Über die Hälfte aller Probanden verhalten sich unter der Faktorstufe „Richtlinie“ illoyal. Dies gilt für das Szenario B (Lounge-Zugang) als auch für D (Meilen-Upgrade), wobei in Szenario D die Neigung zu negativen Verhaltensabsichten stärker ausgeprägt ist. Die wenigsten Nennungen bezüglich der Verhaltensabsichten erfährt die Kategorie „Beschwerde“.

Abbildung 6-7: Negative Verhaltensabsichten in Szenario B und D (Online-Umfrage)



Probanden unter der Faktorstufe „Richtlinie“ äussern vermehrt negative Verhaltensabsichten als Probanden unter der Faktorstufe „Anreiz“ (vgl. Abb. 6-7; A.5.4).

6.5.2 Geschäftsmodell und Fairness-Beurteilung (h8)

1. Formulierung von Hypothese 8

h8 modelltheoretisch formuliert: Zwischen den Gruppen von UV5 (Geschäftsmodell) gibt es signifikante Unterschiede bezüglich der Fairness-Beurteilung (AV3).

2. Stichprobengrösse

Von 1'080 Probanden haben 723 (67%) die Manipulation des Geschäftsmodells verstanden. Bei den Geschäftsmodellen *Low Cost Carrier* und *Network Carrier* lag die Quote der erfolgreichen Manipulationen mit 94 und 69 Prozent höher als bei *Charter Airline* (59%) und *Regional Airline* (50%).

Um die Unklarheiten bei den Geschäftsmodellen *Charter-* und *Regional Airline* zu berücksichtigen, wird h8 mit Probanden getestet, bei denen der Manipulationscheck Geschäftsmodell (UV5) und Kundensteuerung (UV1) funktioniert. In Szenario B beträgt die Stichprobe gemäss diesen Kriterien $n = 302$, in Szenario D ist $n = 220$ Probanden.

3a. Hypothesenprüfung Szenario B (Lounge-Zugang)

Da vier Geschäftsmodelle und somit vier Mittelwerte verglichen werden, wird mit einer ONEWAY ANOVA untersucht (vgl. Tab. 6-15). Probanden, die dem Geschäftsmodell *Low Cost Carrier (LCC)* zugeordnet sind, beurteilen die Fairness verschiedener Kundensteuerungsmassnahmen mit $\bar{x}_{LCC} = 4.01$ ($s = 1.21$) am tiefsten. Bei *Network Carrier (NC)* beträgt der Mittelwert $\bar{x}_{NC} = 4.19$ ($s = 1.12$), während *Charter- (CA)* und *Regional Airline (RA)* mit $\bar{x}_{CA} = 4.32$ ($s = 1.17$) und $\bar{x}_{RA} = 4.40$ ($s = 1.22$) stärker in Richtung „fair“ tendieren. Diese Mittelwerte werden für beide Faktorstufen „Strafe“ und „Anreiz“ erhoben. Die ANOVA zeigt jedoch auf, dass zwischen den Geschäftsmodellen bezüglich der Mittelwerte kein signifikanter Unterschied besteht. Die Nullhypothese kann nicht falsifiziert werden ($p = .17$). Eine weitere Analyse zum Stichprobenmittelwertvergleich mit t -Tests ist nicht nötig.

3b. Hypothesenprüfung Szenario D (Meilen-Upgrade)

In Szenario D (Meilen-Upgrade) sind die Mittelwertsunterschiede zwischen den Geschäftsmodellen geringer ausgeprägt als in Szenario B. Die Nullhypothese wird nicht falsifiziert. ANOVA zeigt mit einem F -Wert von 0.64 und einer Signifikanz von $p = .59$ die deutliche Ablehnung der Alternativhypothese (vgl. Tab. 6-15).

Tabelle 6-15: Auswirkung des Geschäftsmodells auf Fairness-Beurteilung (Online-Umfrage)

h8 (Online): UV5 → AV3	Geschäftsmodell				ANOVA				
	LCC	Network	Charter	Regional	Quadrat- summe	df	Mittel der Quadrate	F	Sig.
B <i>n</i>	94	72	74	62	7.03	3	2.35	1.68	.17
Mittelwert	4.01	4.19	4.32	4.40	415.40	298	1.39		
Stand.abw.	1.21	1.12	1.17	1.22	422.44	301			
D <i>n</i>	71	56	49	44	2.98	3	0.99	0.64	.59
Mittelwert	4.07	4.11	3.88	4.23	332.10	216	1.54		
Stand.abw.	1.21	1.22	1.38	1.16	335.98	219			

Quelle: Eigene Berechnung mit SPSS

4. Effektgrösse und Teststärke

In Szenario B und D wird die Alternativhypothese abgelehnt; die Berechnung der Effektgrösse erübrigt sich.

Die Teststärke beträgt in Szenario B $(1 - \beta) = .67$, in Szenario D ist $(1 - \beta) = .78$. Ein tieferes β in Szenario D bedeutet eine geringere Wahrscheinlichkeit für einen Fehler zweiter Art (irrtümliche Annahme der Nullhypothese).

5. Interpretation und Übersicht (h8)

Die Nullhypothese postuliert, dass sich die vier Stichprobenmittelwerte bezüglich Fairness-Beurteilung nur zufällig unterscheiden. In Szenario B (Lounge-Zugang) und in Szenario D (Meilen-Upgrade) wird die Nullhypothese nicht falsifiziert. Das Geschäftsmodell spielt in der Fairness-Wahrnehmung der Kunden keine Rolle, wenn es um die Bewertung von Kundensteuerungsmassnahmen geht.

6.5.3 Akzeptanzniveau und ex post Image (h9)

1. Formulierung Hypothese 9

h9 modelltheoretisch formuliert: Wenn das Akzeptanzniveau (AV1) tief ist, dann ist die Wahrnehmung des *ex post* Airline Image (AV5) ebenfalls tief.

2. Stichprobengrösse

Für die Bewertung der Auswirkung des Akzeptanzniveaus auf das *ex post Image* werden nur Probanden verwendet, die das „Gesamtmodell“ verstanden haben. Einerseits hängt das Akzeptanzniveau vom Kundensteuerungskonzept ab. Andererseits kann eine adäquate *ex post Image*-Beurteilung nur erfolgen, wenn die Manipulation der Kommunikation verstanden wurde. Es werden nur die Daten

von Probanden verwendet, bei denen die Manipulation bezüglich der Faktoren UV1 (Kundensteuerung) und UV2 (Kommunikation) funktioniert hat. Dieses Auswahlverfahren führt zu $n = 171$ für Szenario B und $n = 133$ für Szenario D.

3a. Hypothesenprüfung Szenario B (Lounge-Zugang)

Zur Prüfung der Hypothese 9 wird auf die Regressionsanalyse zurückgegriffen, da es sich um eine Korrelation zwischen zwei AVn handelt und nicht um Gruppen. Alle untersuchten Variablen sind intervallskaliert.

Die Regressionsanalyse bestätigt, dass der Korrelationskoeffizient in Szenario B mit $r = .13$ tief ausfällt ($n = 171$, $p = .09$). Der r^2 -Wert von $.02$ bedeutet, dass nur 2% der Varianz von der Kriteriumsvariable *ex post Image* durch die Prädiktorvariable „tiefe Akzeptanz“ erklärt werden kann (vgl. Tab. 6-16). Der aus der Variable „tiefe Akzeptanz“ prognostizierte Ausprägungsgrad von *ex post Image* korreliert demnach sehr tief. Mit $p = .09$ ist dieses Ergebnis nicht signifikant und H_0 wird nicht falsifiziert. Dies bedeutet, dass die Variable *ex post Image* in keinem Zusammenhang mit der Variablen „tiefe Akzeptanz“ steht, der nicht auch durch den Zufall erklärt werden kann.

3b. Hypothesenprüfung Szenario D (Meilen-Upgrade)

Die Korrelation mit $r = .22$ erfüllt in Szenario D die Bedingungen (r und df) zur Annahme von H_1 ($n = 133$, $p = .01$). Es darf mit einer geschätzten Irrtumswahrscheinlichkeit von kleiner 1% angenommen werden, dass die untersuchte Stichprobe nicht aus einer Population stammt, in der die interessierenden Merkmalsausprägungen nicht korreliert sind (vgl. Tab. 6-16). Dieses Ergebnis weist eine geringe Varianzaufklärung (r^2) von $.05$ auf. Die Annahme von H_1 muss mit Vorsicht interpretiert werden.

Tabelle 6-16: Auswirkung der Akzeptanz auf (ex post) Image (Online-Umfrage)

h9 (Online): AV1 → AV5		ANOVA				Korrelation		Effekt	
		Quadratsumme	df	Mittel der Quadrate	F	Sig.	r	r ²	d
B	Regression	1.33	1	1.33	2.92	.09	.13	.02	0.26
	Residuen	76.86	169	0.46					
	Gesamt	78.19	170						
D	Regression	2.41	1	2.41	6.37	.01	.22	.05	0.44
	Residuen	49.66	131	0.38					
	Gesamt	52.08	132						

Quelle: Eigene Berechnung mit SPSS und G*Power

4. Effektgrösse und Teststärke

Da in Szenario B die Alternativhypothese abgelehnt wird, kann aufgrund der Effektgrösse keine Aussage gemacht werden. Die Teststärke, d.h. die Wahrscheinlichkeit, dass H_0 korrekterweise abgelehnt wird, beträgt in Szenario B $(1 - \beta) = .64$.

In Szenario D ergibt das Differenzmass einen mittleren Effekt von $d = .44$. Die Teststärke ist mit einem Wert von $(1 - \beta) = .92$ hoch.

5. Interpretation und Übersicht (h9)

Die Alternativhypothese wird in Szenario B (Lounge-Zugang) knapp abgelehnt und in Szenario D (Meilen-Upgrade) mit einer tiefen Varianzaufklärung angenommen.

Eine tiefe Akzeptanz einer Kundensteuerung hat nicht zwangsläufig eine negative Auswirkung auf das *ex post* Airline Image. Das Antwortverhalten der Probanden unterscheidet sich je nach Szenario. Die Auswirkung eines nicht gewährten Meilen-Upgrades ist für die *ex post* Wahrnehmung des Airline Image gravierender als ein fehlender „Lounge-Zugang“ mit „VIP Boarding“. Diese Tatsache hängt damit zusammen, dass in Szenario B die Anreiz-Option bestehen bleibt, während in Szenario „Meilen-Upgrade“ die Probanden vor eine unumgängliche Tatsache gestellt werden (Richtlinie ohne Option, d.h. „Bestrafung“). Da in der vorliegenden Arbeit das Bestrafungskonstrukt im Vordergrund steht und damit Szenario D, wird die Variable *ex post Image* im Modell bestätigt.

6.5.4 Fairness-Beurteilung und Akzeptanzniveau (h10)

1. Formulierung Hypothese 10

h10 modelltheoretisch formuliert: Wenn die Fairness-Beurteilung (AV3) hoch ist, dann ist die Akzeptanz (AV1) ebenfalls hoch.

2. Stichprobengrösse

Wie bei h9 werden unter h10 nur diejenigen Probanden untersucht, die das Gesamtkonzept verstanden haben. Diese strenge Selektion erfolgt, weil h10 eine neu untersuchte Wirkbeziehung im Forschungsmodell betrifft. Die Gruppe „erfolgreiche Manipulation Kundensteuerung und Kommunikation“ beträgt $n = 171$ in Szenario B und $n = 133$ in Szenario D (vgl. Kap. 6.5.3).

3a. Hypothesenprüfung Szenario B (Lounge-Zugang)

Die Modellzusammenfassung zeigt eine Spearman-Korrelation von $r = .22$ ($n = 171$, $p = .00$) zwischen der AV3 (Fairness) und AV1 (Akzeptanz). Die Alternativhypothese von h10 wird gemäss *F*-Test auf dem Signifikanzniveau $p < 0.1\%$ angenommen, sie erfüllt die Anforderung von $r = .21$ bei $df = 150$ (vgl. A.3.1). Die Varianzerklärung beträgt 5%.

3b. Hypothesenprüfung Szenario D (Meilen-Upgrade)

In Szenario D ist die direkte Korrelation zwischen Fairness und Akzeptanz (ohne *emotionale Reaktion*) mit einer Varianzaufklärung von 11% und einer Korrelation von $r = .33$ ($n = 133$, $p = .00$) deutlich höher (vgl. Tab. 6-17). Die Alternativhypothese wird angenommen (vgl. A.3.1 zu den verwendeten Mindestgrößen). Bei der Interpretation dieser Korrelation kann der Zufall mit einer Irrtumswahrscheinlichkeit von $p < 0.01\%$ ausgeschlossen werden.

Tabelle 6-17: Zusammenhang zwischen Fairness und Akzeptanz (Online-Umfrage)

h10 (Online): AV3 → AV1		ANOVA				Korrelation		Effekt	
		Quadrat- summe	df	Mittel der Quadrate	F	Sig.	r	r ²	d
B	Regression	24.64	1	24.64	8.34	.00	.22	.05	0.44
	Residuen	502.22	169	2.95					
	Gesamt	526.86	170						
D	Regression	54.72	1	54.72	20.43	.00	.33	.11	0.70
	Residuen	447.29	131	2.68					
	Gesamt	502.00	132						

Quelle: Eigene Berechnung mit SPSS und G*Power

4. Effektgröße und Teststärke

Trotz einer tiefen Korrelation in Szenario B wird mit $d = 0.44$ ein mittelstarker Effekt nachgewiesen. Die Teststärke in Szenario B ist $(1 - \beta) = 0.59$.

In Szenario D liegt eine mittlere Ausprägung der Effektgröße von $d = 0.70$ vor. Die Wahrscheinlichkeit, dass H_0 korrekterweise abgelehnt wird, ist im Vergleich zu B höher mit $(1 - \beta) = .77$.

5. Interpretation und Übersicht (h10)

In beiden Szenarien B (Lounge-Zugang) und D (Meilen-Upgrade) wird die Alternativhypothese angenommen. Je höher die Fairness-Beurteilung, umso höher ist das Akzeptanzniveau der Probanden. Dieser Zusammenhang ist in Szenario D stärker ausgeprägt mit einer Effektgröße von $d = 0.70$. Hypothese 10 hat bestätigt, dass die Variable *Fairness-Beurteilung* (als unabhängige Variable) das Akzeptanzniveau eines Sachverhalts genauso gut wie die Variable *emotionale Reaktion* vorhersagt.

6.6 Empirische Grenzen

Wie jede Forschungsarbeit hat diese Untersuchung empirische und methodische Grenzen. Diese Grenzen gelten für die Flughafen- und die Online-Umfrage, da in beiden Untersuchungen experimentell, basierend auf dem zentralen (angepassten) Forschungsmodell, gearbeitet wurde.

6.6.1 Empirisch bedingte Grenzen

Grenzen bei den experimentellen Bedingungen

Mit fünf experimentellen Bedingungen war aus statistischen Überlegungen eine umfangreiche Stichprobengrösse nötig (vgl. Kap. 5.1.6). Die Abstufung der Bedingungen wurde daher aus forschungsökonomischen Gründen nur mit zwei Faktoren durchgeführt (ausser bei UV5). Die Verwendung der Stufen „hoch“ vs. „tief“ führte bei den Probanden zu einem grossen Interpretationsspielraum. Die „feinere“ Abstufung der Faktoren könnte bei weniger untersuchten Faktoren zu differenzierteren Ergebnissen führen, bei gleichzeitig hoher Teststärke. Eine differenzierte Ausprägung der Faktorstufen würde die externe Validität erhöhen.

Grenzen bei den abhängigen Variablen

Die verwendeten Skalen stützen sich auf frühere Untersuchungen und auf Ergebnisse der qualitativen Empirie. Potentielle Verbesserungen betreffen insbesondere die Variable *emotionale Reaktion*, wie die Beiträge verschiedener Disziplinen im theoretischen Teil aufzeigen. Die Messung der emotionalen Reaktion sowie der Fairness-Beurteilung mit einer 7er Rating-Skala, insbesondere bei einem fiktiven Szenario, stösst an die Grenzen der Diskriminierungsfähigkeit von Probanden. Zur Erhöhung der Reliabilität könnte eine vereinfachte Skala mit weniger Ausprägungen eingeführt und getestet werden. Bspw. konnte im Online-Fragebogen durch die Anpassung des Manipulationschecks UV1 bei gleichzeitiger Verwendung einer 3er anstatt 6er Likert-Skala die Manipulation einfacher und für die Probanden verständlicher durchgeführt werden.

Reaktanz entstand bei Probanden vor allem unter der Faktorstufe „Anreiz“ bei den neg. Verhaltensabsichten, wie einzelne Kommentare zeigen. Da aus methodischen Gründen versucht wurde, ein Szenario sowohl in der Anreiz- als auch in der Strafsituation möglichst äquivalent darzustellen, waren die Skalen zu den negativen Verhaltensabsichten für einzelne Probanden irreführend. Sie wurden jedoch nicht in die Auswertung mit einbezogen.

Grenzen der Stichprobe

Bei der Auswahl der Probanden ergeben sich zwangsläufig Verzerrungen bei einer Stichprobe. Um eine hohe statistische Validität zu erreichen, sollte als wirkungsvolles Verfahren die Zufallsauswahl eingesetzt werden (Sarris, 1999). In keinem der beiden vorliegenden Experimente konnte die Auswahl völlig zufällig durchgeführt werden. Indem bei der Flughafen-Umfrage Probanden

an verschiedenen Tagen und verschiedenen Gates ausgewählt wurden, wurde versucht, dieser Anforderung nachzukommen. Bei der Online-Umfrage wurde durch die Auswahl mit verschiedenen Newsgroups und E-Mail-Listen eine möglichst breite Streuung angestrebt. Trotz Erweiterung der Stichprobenszusammensetzung durch die Online-Untersuchung zeigen die demografischen Variablen, dass bspw. ein grosser Teil der Probanden in beiden Experimenten aus der Schweiz stammt und mit Swiss fliegt. Die Replikation der Flughafen-Umfrage an anderen internationalen Flughäfen würde diese Verzerrungseffekte vermindern (Erhöhen der inferenzstatistischen Validität). Die geforderte Stichprobengrösse und das Allokationsverhältnis zwischen n_1 und n_2 werden bei allen Hypothesen erfüllt (vgl. Kap. 5.1.6). Die einzige Ausnahme sind die demografischen Hypothesen zur Klassenzugehörigkeit, wo die Gruppe *Business Class* leicht untervertreten ist.

Grenzen bezüglich der internen und externen Validität

Die Störvariablen und das Design eines Experiments beeinflussen die interne Validität. In dieser Arbeit konnte die interne Validität mit der Wahl eines *Between-Subject-Designs* und eines „Quasi“-Feldexperiments erhöht werden. Trotzdem gibt es Einschränkungen bei der internen Validität. Trotz eines Manipulationschecks und eines Pretests zeigen die hohen Ausfallquoten („Manipulationscheck nicht bestanden“), dass v.a. in der Online-Untersuchung entweder nicht sorgfältig gelesen oder die Manipulationen nicht verstanden wurden. Zudem können interaktive Effekte (Probanden wurden unter verschiedenen Bedingungen untersucht) zu einer tiefen Validität geführt haben. Eine konsequentere Umsetzung des Max-Kon-Min-Prinzips (vgl. Kap. 5.1.1) würde die interne Validität erhöhen.

Mit der Szenario-Technik werden die Probanden gebeten, sich „virtuell“ in die beschriebene Situation hineinzusetzen. Mit dieser Vorgehensweise können viele Situationen und Variablen untersucht werden. Die Methode hängt aber stark von der Diskriminierungsfähigkeit der Probanden ab. Besser geeignet, aber sehr aufwendig, wäre eine Beobachtung bei Probanden, welche soeben „bestraft“ wurden. Eine reale Manipulation mit einem Experiment durch „Bestrafung“ am Flughafen und anschliessende Messung der Reaktion ist aus forschungsethischen Gründen nicht möglich.

Mit der Versuchsanordnung am Flughafen konnte die externe Validität und somit die Situationsrepräsentativität gewährleistet werden. Die Forschungsergebnisse können von der Stichprobe auf die Population „Passagier“ übertragen werden, weil die Hypothesen und das Modell in dieser Arbeit durch theoretische Aufarbeitung und weitere Forschungsarbeiten gestützt sind. Ob sich die Probanden aber in der realen Umgebung gleich verhalten wie im Fragebogen beschrieben, ist fraglich.

6.6.2 Methodisch bedingte Grenzen

Die Auswertung mit einer statistischen Methode hat Vor- und Nachteile (vgl. A.3). Die in dieser Arbeit verwendeten Verfahren werden ausgewählt, weil sie die Forschungsziele unterstützen. Durch

die Kombination qualitativer mit quantitativen Methoden sowie die Versuchsanordnung mit zwei Experimenten wurde versucht, die methodischen Problemfelder in engen Grenzen zu halten.

Grenzen bezüglich der Modellkausalität

Die zur Modellevaluation verwendeten Kriterien und statistischen Verfahren deuten darauf hin, dass die empirisch erhobenen Daten mit dem Forschungsmodell übereinstimmen. Der Zufall kann bei Ablehnung der Nullhypothesen zur Erklärung einer Ursache auf einem bestimmten Signifikanzniveau ausgeschlossen werden (Sarris, 1999; Kirk, 1995). Weshalb ein Unterschied besteht, wird aber mit entscheidungsstatistischen Verfahren nicht belegt. Um dies zu klären, wird Rückgriff genommen auf inhaltliche Interpretationen im theoretischen A-priori-Forschungsmodell. Diese wiederum hängen von der Interpretationsfähigkeit des Forschers ab. Die Konsistenz der analysierten Modelle muss nicht unbedingt eine Konsistenz mit der Realität implizieren (Bollen, 1989).

Auch von einem hohen Korrelationswert allein kann *formal* nicht auf eine kausale Beziehung zwischen den Merkmalen geschlossen werden. Ein hoher Korrelationswert ist eine notwendige, aber keine hinreichende Bedingung für eine kausale Beziehung zwischen zwei Merkmalen (Cramer und Kamps, 2007; Schira, 2003). Die hinreichende Bedingung zur Kausalität wird durch das theoretisch-konzeptionelle Forschungsmodell postuliert.

Grenzen bezüglich der Verteilungseigenschaften

Die in jeder Hypothese untersuchten abhängigen Variablen sollten bei prüfstatistischen Verfahren näherungsweise normalverteilt sein (vgl. A.3.2). Für alle Hypothesen in der vorliegenden Arbeit, bei denen die Stichprobe gross ist ($n > 100$), wurde die Überprüfung der Normalverteilung nicht mit dem *Kolmogorov-Smirnov-Test*, sondern mit Q-Q Plots durchgeführt (vgl. A.3.2).

Die grafische Auswertung jeder untersuchten abhängigen Variablen in beiden Experimenten (AV1 - AV5) zeigt auf, dass in der vorliegenden Untersuchung von näherungsweise normalverteilten Variablen ausgegangen werden kann. Eine Ausnahme bildet AV2 (neg. Verhaltensabsichten), welche bei der grafischen Analyse leichte Ausreisser zeigt. Diese Ausreisser werden jedoch durch die Verwendung einer grossen Stichprobe ($n = 275$) relativiert und es kann angenommen werden, dass sie ebenfalls normalverteilt sind (vgl. Hirsig, 2001).

Der *Kolmogorov-Smirnov-Test* hat bei Hypothesen mit kleiner Stichprobengrösse (h_2, h_{d1}) ebenfalls näherungsweise normalverteilte Variablen bestätigt.

6.7 Überprüftes Forschungsmodell

Das provisorische Forschungsmodell für das erste Experiment wird mit der Flughafen-Untersuchung bestätigt. Drei experimentelle Bedingungen beeinflussen die Fairness-Beurteilung der Probanden gleichsinnig (positiv):

- Den stärksten Effekt hat die *Kundensteuerung* in Szenario C (Rückreisedatum) und D (Meilen-Upgrade): Die Fairness-Beurteilung der Probanden hängt von der Gestaltung des Kundensteuerungskonzepts ab. Passagiere beurteilen ein identisches Szenario als „fairer“, wenn es als Anreiz anstatt als Richtlinie gestaltet wird.
- Ebenfalls stark ausgeprägt ist die experimentelle Bedingung *ex ante Image* in den Szenarien B (Lounge-Zugang) und D (Meilen-Upgrade). Je höher das Image einer Airline vor einem Ereignis eingestuft wird, umso loyaler verhalten sich Probanden. Das *ex ante Image* hat keinen moderierenden Effekt auf die Kundensteuerung oder Kommunikation. Sie wirkt unabhängig auf die Fairness-Beurteilung.
- Die verständliche und deutliche *Kommunikation* der Anreiz-Optionen und Bedingungen eines Tickets trägt in allen Szenarien zur höheren Fairness-Beurteilung bei.

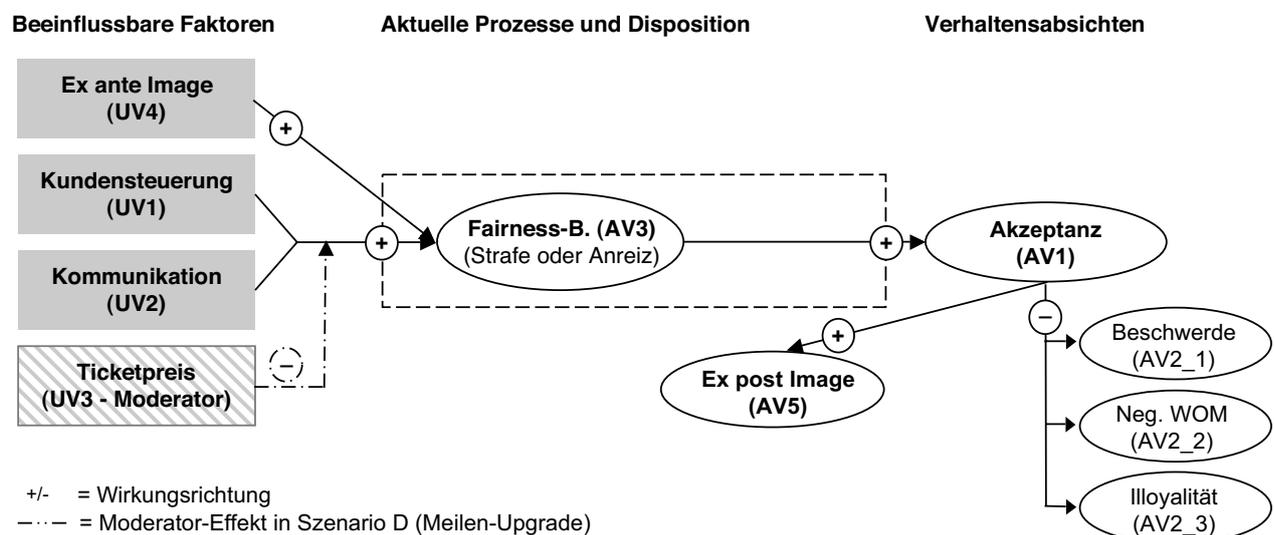
Der Preis eines Tickets hat keinen Einfluss auf die Fairness-Beurteilung, aber in Szenario D (Meilen-Upgrade) einen leicht gegensinnigen (negativen) Interaktionseffekt auf *Kundensteuerung* und *Kommunikation* (vgl. Abb. 6-8).

Die Fairness-Beurteilung entscheidet über das Akzeptanzniveau, welches in allen Szenarien gegensinnig mit den negativen Verhaltensabsichten korreliert. Die Neigung zu negativen Verhaltensabsichten ist unter der Faktorstufe „Richtlinie“ ausgeprägter als unter „Anreiz“. Als ausgeprägteste negative Verhaltensabsicht wird „Illoyalität“ identifiziert. Somit korrelieren die Fairness-Beurteilung, Akzeptanz und Loyalität positiv gleichsinnig miteinander. Diese Aussage wird von der Versuchsanordnung gestützt, denn in den Szenarien A, B und C wurden keine „harten“ Bestrafungen simuliert, sondern den Probanden jeweils die Möglichkeit offen gelassen, die Option nachträglich zu „kaufen“. Die endogenen demografischen Variablen *Flugerfahrung Privatflüge* und *Klassenzugehörigkeit* bestätigen in allen Szenarien ihren Einfluss auf das Akzeptanzniveau. Erfahrene Privatreisende und *Business-Class*-Passagiere weisen bei einem negativen Ereignis eine höhere Akzeptanz auf als unerfahrene Privatreisende und *Economy-Class*-Passagiere.

In den Szenarien A (Übergepäck) und C (Rückreisedatum) besteht eine Diskrepanz zwischen dem Verhalten (Zahlen der Gebühr / Akzeptanz) und der Einstellung (Fairness-Beurteilung). Obwohl bei der Steuerung über Anreize die Fairness-Beurteilung hoch ist, ist die tatsächliche Nachfrage nach Online-Zusatzdienstleistungen in diesen Szenarien tief.

Im zweiten Experiment werden die Ergebnisse der Flughafen-Untersuchung bezüglich der zentralen Wirkung der Hauptvariablen *Kundensteuerung*, *Kommunikation* und *Akzeptanz* mit der Online-Untersuchung und einer erweiterten Population bestätigt. Der neu untersuchte Zusammenhang zwischen Fairness-Beurteilung und Akzeptanz (ohne Variable *emotionale Reaktion*) wird ebenfalls bestätigt. Die Auswirkung der Akzeptanz auf das *ex post Image* ist abhängig vom untersuchten Szenario: der Einfluss ist in Szenario D (Meilen-Upgrade) gleichsinnig ausgeprägt, in B (Lounge-Zugang) wird die Hypothese knapp abgelehnt. Bei der Bewertung verschiedener Kundensteuerungsmassnahmen hat das Geschäftsmodell einer Airline keinen Einfluss auf die Fairness-Beurteilung. Mit dem empirisch überprüften Wirkungsmodell in Abbildung 6-8 wird die Forschungsfrage c.1 beantwortet und damit ein Beitrag an die Wissenschaft geleistet.

Abbildung 6-8: Forschungsmodell nach der Überprüfung



Quelle: Eigene Darstellung

Im Wirkungsmodell sind die Variablen emotionale Reaktion und Geschäftsmodell nicht aufgeführt. In der Kundenwahrnehmung sagt die Fairness-Beurteilung die Akzeptanz einer Kundensteuerungsmassnahme genauso vorher wie die emotionale Reaktion. Das Geschäftsmodell und der Ticketpreis als experimentelle Bedingungen bleiben gegenüber der Fairness-Beurteilung wirkungslos. Abbildung 6-8 zeigt diesen Sachverhalt, indem die überprüften Zusammenhänge angepasst und mit der Wirkungsrichtung aufgeführt werden. Der moderierende Einfluss vom Ticketpreis in Szenario D (Meilen-Upgrade) ist mit gestrichelter Linie dargestellt.

7 Implikationen für Forschung und Praxis

Kundenverhalten während des Dienstleistungsprozesses kann für Unternehmen schädigend sein. Dienstleistungsanbieter, insbesondere Airlines, versuchen aufgrund hoher Fixkosten den Kunden mit Bestrafungsrichtlinien zu steuern, damit er sich leistungskonform verhält. Damit werden zusätzliche Kosten oder entgangene Gewinne vermieden. Mit diesem Kapitel wird aufgezeigt, wie unerwünschtes Kundenverhalten verhindert und mit alternativen Steuerungsmassnahmen zusätzlich Umsatz generiert werden kann.

Mit den *theoretisch-konzeptionellen Grundlagen* wurden in dieser Arbeit die bestehende Forschung zur Bestrafung disziplinenübergreifend betrachtet, Begriffe und Konstrukte abgegrenzt sowie das provisorische Wirkungsmodell erarbeitet.

Im *qualitativen Teil* der Arbeit wurden Kundensteuerungsmassnahmen im Airline-Bereich untersucht. Um das aus der Forschungsperspektive zentrale Konstrukt *Kundenstrafen* erfassen zu können, wurden eine Umfrage und eine Beobachtung von Passagieren am Check-In durchgeführt (Kundenperspektive). Einsichten aus der Anbieterperspektive halfen, alternative Massnahmen zur Bestrafung zu entwickeln. Diese Massnahmen basieren auf verschiedenen Szenarien, die eine experimentelle Untersuchung ermöglichen. Das provisorische Wirkungsmodell wurde zum zentralen A-priori-Forschungsmodell erweitert.

Im *quantitativen Teil* wurden wichtige Einflussfaktoren auf die Kundenwahrnehmung und das Kundenverhalten, basierend auf dem zentralen A-priori-Forschungsmodell, experimentell getestet. Dabei wurden zwei verschiedene Erhebungsmethoden eingesetzt, um die Güte der vorliegenden Untersuchung zu erhöhen. Ergebnisse der statistischen Datenauswertung zeigen, dass Massnahmen zur Kundensteuerung von Kunden bezüglich der *Fairness* verschieden beurteilt werden. Bei einer tiefen Fairness-Beurteilung wird eine Massnahme als Strafe wahrgenommen. Es resultiert ein tiefes Akzeptanzniveau. Kunden, die sich bestraft fühlen (und Strafe als unfair einstufen), reagieren mit negativen Verhaltensabsichten, insbesondere mit illoyalem Verhalten. Zudem wird die *ex post* Image-Wahrnehmung nach einer Bestrafung negativ beeinflusst. Negative Verhaltensabsichten haben Auswirkungen auf das Unternehmensergebnis, und zwar höhere als der Nutzensgewinn konformen Verhaltens. Werden anstatt Strafen *Anreiz-Optionen* eingesetzt, führt die Kundensteuerung nicht zu negativen Verhaltensabsichten. Der Kundenvorteil und damit der Kundenwert werden erhöht.

In Anlehnung an die Struktur der Forschungsfragen aus Kapitel 1.4.2 werden in Kapitel 7.1 die Forschungsfragen a.1 und a.2 aus theoretischer Perspektive beantwortet. Die Implikationen aus der empirischen Untersuchung werden mit den Forschungsfragen b.1, b.2 und b.3 in Kapitel 7.2 beantwortet. Die konkrete Umsetzung der Implikationen auf das Dienstleistungs- und das Airline-Management ist in Kapitel 7.3 erläutert (c.2). Bevor in Kapitel 7.5 der weitere Forschungsbedarf dargestellt wird, werden die Ergebnisse in Kapitel 7.4 kritisch gewürdigt.

7.1 Implikationen aus dem theoretischen Teil

Die Forschungsfrage a.1 betrifft die theoretischen Erkenntnisse zur individuellen Wahrnehmung von Strafen und der darauf basierenden Verhaltensreaktion. Die Kundenwahrnehmung als Fairness-Beurteilung wird gemäss theoretischer Erkenntnisse beeinflusst von (vgl. Kap. 7.1.1):

- der Kundensteuerungsmassnahme (Push- oder Pull) und deren *Framing* (Kommunikation);
- der *ex ante* Image-Wahrnehmung (Interaktion mit Kommunikation und Kundensteuerung);
- Equity-Überlegungen. Der Ticketpreis und die dem Unternehmen verursachten Kosten werden gegenübergestellt;
- demografischen Variablen wie der Klassenzugehörigkeit und der Produkterfahrung;
- der Wahrscheinlichkeit, erwischt zu werden (bei Bestrafungsrichtlinien), und von Drittpersonen sowie der Atmosphäre vor Ort.

Die Verhaltensreaktion von Kunden basiert gemäss theoretischer Erkenntnisse auf (vgl. Kap. 7.1.2):

- der emotionalen Reaktion, die durch die kognitive Fairness-Wahrnehmung einer Massnahme bestimmt wird;
- der kognitiven Beurteilung der Fairness einer Massnahme. Eine Strafe führt zu einer tiefen Fairness-Beurteilung, eine Anreiz-Option zu einer hohen Fairness-Beurteilung. Bei einer tiefen Fairness-Beurteilung einer Massnahme folgen negative Verhaltensabsichten (Illoyalität, neg. WOM und Beschwerden) und eine negative Gesamtbeurteilung des Unternehmens.

Der mit der Forschungsfrage a.2 geforderte Beitrag verschiedener Disziplinen zur Bestrafungsthematik wird in Kapitel 7.1.1 bei der Beantwortung der Forschungsfrage a.1 berücksichtigt. Aus der Synthese zwischen der Anbieter- und der Kundenperspektive lassen sich Handlungsempfehlungen ableiten. Die theoretischen Beiträge aus Sicht der Jurisprudenz, der Psychologie, der Ökonomie und der Soziologie sind in Kapitel 2.2 bis 2.4 detailliert aufgeführt.

7.1.1 Wahrnehmung der Bestrafungssituation

Aus Sicht des Kunden muss für eine Strafe eine Schuld bestehen - ohne Schuld keine Strafe (vgl. Kap. 2.2). „Schuld“ bedeutet im Kontext der Dienstleistungserstellung, dass Kunden Einkaufsvereinbarungen missachten. Im Gegensatz zu kritischen Ereignissen geht es bei der Bestrafung um ein Fehlverhalten von Kunden (vgl. Kap. 2.1.3). Aus Unternehmenssicht ist die hauptsächliche Wirkung von Strafen die Kompensation und die Prävention (vgl. Kap. 2.1.2).

Kompensation

Aus der psychologischen Equity-(Fairness-) Wahrnehmung folgt, dass bei der Festlegung der Strafhöhe das Verhältnis zwischen Input und Outcome berücksichtigt werden muss (vgl. Kap. 2.3.2). Passagiere bspw. verstehen, dass sie Übergepäckgebühr bezahlen müssen. Die Kosten für Übergepäck sollten aber in einem sinnvollen Verhältnis zum Ticketpreis sein. Ansonsten empfinden Kunden die Strafzahlung als rein gewinnorientiertes Handeln der Unternehmung (Kompensation). Dies führt zu einer tiefen Akzeptanz und damit zu negativen Verhaltensabsichten.

Anwesende Drittpersonen und die soziale Atmosphäre vor Ort beeinflussen das Verhalten von Kunden. Prinzipien der *Equality*-Wahrnehmung helfen bei der Ausgestaltung der Kundensteuerung in einer Community. Gemäss der Equality-Wahrnehmung sollen alle Kunden gleich behandelt werden (vgl. Kap. 2.3.2). Dies erweist sich im Dienstleistungskontext als schwierig. Kunden mit einem hohen Status im Bonusprogramm einer Airline bspw. erwarten Rücksicht bei Bestrafungen (vgl. Kap. 3.1.4). Um die Equality bei Passagieren in einem solchen Fall nicht zu beeinträchtigen, sollten Kunden, die sich nicht leistungskonform verhalten und denen eine Strafe erlassen wird, getrennt von den Mit-Kunden instruiert werden (vgl. Kap. 2.4.3). Falls man einzelnen Kundensegmenten im Sinne einer Wiedergutmachung entgegenkommen will, sollten Strafen nicht teilweise, sondern ganz erlassen werden. Eine Strafe zu vermindern, bedeutet ein Schuldeingeständnis für falsches Verhalten von Seiten der Unternehmung (vgl. Kap. 3.1.7). Dem Kunden muss jedoch klargemacht werden, dass er sich nicht leistungskonform verhalten hat.

Prävention

Konsumenten sind bis zu einer gewissen Strafhöhe indifferent. Strafen ziehen nicht zwangsläufig negative Verhaltensabsichten nach sich, ausser es findet ein *over-punishing* statt. Die Strafhöhe sollte gemäss ökonomischen Theorien der Normverletzung angepasst werden (vgl. Kap. 2.4.1). Dies erfordert jedoch geschulte Mitarbeiter mit Entscheidungskompetenz, flexible Prozesse und Handlungsanleitungen zum Verhalten bei kritischen Situationen. Bei Strafmassnahmen spielt aus ökonomischer Sicht die Wahrscheinlichkeit, erwischt zu werden, eine grosse Rolle. Grundsätzlich unterstützen Strafen leistungskonformes Verhalten. Nur auf die Legalität einer Kundenhandlung zu achten, genügt aber nicht. Selbst wenn alle anderen Attribute der Dienstleistungserstellung wunschgemäß erfüllt wurden, führt eine Strafe wegen Vertragsverletzung zu einer negativen Gesamtbeurteilung des Unternehmens (vgl. Kap. 2.3.4). Es spielt dabei keine Rolle, ob es sich um monetäre oder nicht monetäre Strafen handelt.

Aus der Organisationssoziologie folgt, dass das von Konsumenten empfundene Machtgefälle zwischen Organisationen und Konsumenten bei deutlicher Kommunikation der Steuerungsmassnahmen verhindert wird (vgl. Kap. 2.4.2). Konsumenten fragen sich bei einer Bestrafung, welchen Schaden sie der Organisation beigefügt haben, und vergleichen die vermuteten Kosten des Schadens mit der Strafgebühr. Stimmt dieses Verhältnis nicht, folgen negative Verhaltensabsichten. Eine klare Dar-

stellung verursachter Kosten nicht leistungskonformen Verhaltens trägt dazu bei, bei Konsumenten Verständnis für die präventiven Massnahmen zu entwickeln.

7.1.2 Verhaltensabsichten

Die Verhaltensabsichten können als Prediktor für tatsächliches Kundenverhalten eingesetzt werden (vgl. Kap. 2.1.4). Sie werden in dieser Arbeit, im Kontext von Bestrafungssituationen, gegliedert als: Beschwerde einreichen, negative Word-of-Mouth (Mund-zu-Mund-Propaganda) und Illoyalität (kein Wiederkauf der Leistung oder Wechsel des Anbieters). Diese Absichten können einzeln oder kombiniert auftreten. Sie werden vom Konstrukt *emotionale Reaktion* ausgelöst (vgl. Kap. 2.5.3).

Bei *High-Involvement*-Produkten wie einer Flugreise ist die Neigung zur Beschwerdeeinreichung höher als bei *Low-Involvement*-Produkten (vgl. Kap. 3.1.5). Forschung zum Beschwerdeverhalten zeigt, dass Kunden, die sich beschwerten, am loyalsten sind (vgl. Kap. 2.1.4).

Negative WOM von Kunden darf von Firmen nicht unterschätzt werden; immerhin werden zwei Drittel aller Verkäufe im Konsumgüterbereich über WOM getätigt (Solomon, 2006). Es wird angenommen, dass im Airline-Bereich WOM genauso entscheidend ist, da es sich bei einer Flugreise um ein „Vertrauensgut“ handelt (vgl. Kap. 3.1.1).

Wiederholungskäufe bedeuten nicht zwangsläufig eine enge Bindung an eine Firma. Selbst wenn Kunden zufrieden mit einer Dienstleistung sind, wechseln sie den Anbieter (vgl. Kap. 3.1.5). Bei *Low Cost Carrier* bspw. steht nicht das Image, sondern der Preis und das Angebot im Vordergrund. Sobald diese Produktattribute nicht mehr stimmen, wechseln Kunden den Anbieter (vgl. Kap. 3.1.2).

Fühlen sich Kunden in ihren Rechten und Ansprüchen verletzt oder fällt die Equity-Wahrnehmung negativ aus, können neben Illoyalität, negative WOM oder Beschwerden auch Drittparteien (Medien) involviert oder kann rechtlich gegen das Unternehmen vorgegangen werden. Neben einem Imageschaden entstehen damit auch monetäre sowie nicht monetäre Kosten für eine Organisation (vgl. Kap. 2.1.1).

7.2 Implikationen aus dem empirischen Teil

In diesem Kapitel werden die empirischen Ergebnisse zur Beantwortung der Forschungsfragen b.1, b.2 und b.3 aus der Airline-Perspektive interpretiert (vgl. Kap. 1.4.2).

Die erste Forschungsfrage (b.1) betrifft die Einflussfaktoren auf die Kundenwahrnehmung verschiedener Kundensegmente. Die empirisch bestätigten Einflussfaktoren sind (vgl. Kap. 7.2.1; 7.3.2):

- die Art der Kundensteuerung (Bestrafungsrichtlinie oder Anreiz-Option) und deren Kommunikation (klar oder unklar). Diese Faktoren wirken gleichsinnig (positiv) auf die Fairness-Beurteilung und auf die *ex post* Image-Wahrnehmung;

- das *ex ante* Image einer Airline. Je höher die Image-Wahrnehmung, umso höher ist die Fairness-Beurteilung: Strafen werden als weniger unfair beurteilt. Das Image hat keinen moderierenden Effekt auf die Variablen Kundensteuerung und Kommunikation.
- Es konnte kein direkter Einfluss des Ticketpreises auf die Fairness-Beurteilung nachgewiesen werden, aber ein Interaktionseffekt auf die Kundensteuerung und auf die Kommunikation. Das Geschäftsmodell hat auch keinen Einfluss auf die Fairness-Beurteilung.
- Die Klassenzugehörigkeit (*Economy* oder *Business Class*) hat einen gleichsinnigen Einfluss auf die Akzeptanz, aber keinen Einfluss auf die Verhaltensabsichten. Die Flugerfahrung hat je nach Reiseart (geschäftlich oder privat) einen (marginalen) Einfluss auf die Akzeptanz und die Verhaltensabsichten.

Die zweite Forschungsfrage (b.2) betrifft die Service-Stufen der Kundensteuerung in der Airline-Branche: Airlines steuern Flugreisende mit Richtlinien, die oft unklar während des Buchungsprozesses in den Tarifbedingungen aufgeführt sind. Dies führt dazu, dass Flugpassagiere am Check-In unerwartet „bestraft“ werden. Einzelne Airlines haben begonnen, anstatt Richtlinien Anreiz-Optionen während des Internet-Buchungsprozesses zu offerieren (vgl. Kap. 7.2.2).

Die dritte Forschungsfrage (b.3) beinhaltet eine erste Implikation zur Umsetzung von Kundensteuerungsmassnahmen für leistungskonformes Verhalten ohne Bestrafung: Verschiedene über den Internetkanal angebotene Optionen eignen sich zur effizienten Kundensteuerung, ohne negative Verhaltensabsichten auszulösen (vgl. Kap. 7.2.3). Im Airline-Bereich kann das Bonusprogramm zur Kundensteuerung über Anreize eingesetzt werden.

7.2.1 Einflussfaktoren auf Wahrnehmung und Verhaltensabsichten

Die Kundenwahrnehmung hängt von verschiedenen Faktoren ab und ist individuell von Erfahrungen und dem Vorwissen geprägt (vgl. Kap. 6.3.8). Trotzdem lassen sich anhand der Ergebnisse der vorliegenden Untersuchung relevante von nicht relevanten Einflussfaktoren auf das Kundenverhalten unterscheiden. Für Airlines werden je nach Situation verschiedene Massnahmen empfohlen (vgl. Kap. 7.3.2).

Die Fairness-Beurteilung einer Anreiz-Massnahme hängt wesentlich von der Preisgestaltung der Anreiz-Option ab (vgl. Kap. 7.3.2). Das überprüfte Forschungsmodell bestätigt, dass die richtig kommunizierte Kundensteuerung über Anreize zu einer hohen Fairness-Beurteilung führt, diese wiederum zu einer hohen Akzeptanz und eine hohe Akzeptanz zu loyalen Verhalten (vgl. Kap. 6.7). Sich ausschliesslich auf die Fairness-Beurteilung zu verlassen, führt aber nicht zur optimalen Kundensteuerung. Nicht alle Massnahmen, die zu einer hohen Fairness-Beurteilung führen, werden tatsächlich nachgefragt (vgl. Kap. 7.3.2).

Zur Kommunikation von Kundensteuerungsmassnahmen eignet sich der Distributionskanal Internet. Die experimentelle Untersuchung hat aufgezeigt, dass eine verständliche und deutliche Kommunikation von Kauf-Optionen oder der (Tarif-) Bedingungen entscheidend zur Fairness-Wahrnehmung beiträgt (vgl. Kap. 6.3.2). Durch die Zunahme an Buchungen über Airline Websites und der damit verbundenen, erschwerten Kommunikation wird sich dieses Problem verschärfen.

Der Flugpassagier hat sich daran gewöhnt, dass die Preise der Flugtickets je nach Buchungszeitpunkt und -klasse ändern (vgl. Kap. 3.1.2). Aus Sicht einer effizienten Kundensteuerung sollten aber ständig ändernde Optionen und Richtlinien vermieden werden, da sonst Überforderung eintritt. Zudem können Optionen und Richtlinien auf dem Internet leichter zwischen Airlines verglichen werden, was die Preissensitivität von Kunden erhöht (vgl. Kap. 3.1.3). Die qualitative Inhaltsanalyse hat gezeigt, dass der Wert eines Tickets (*Ticketvalue*) auf den untersuchten Airline-Buchungsseiten unvollständig und undeutlich kommuniziert wird (vgl. A.5.2). Unter *Ticketvalue* wird der Leistungsumfang eines Flugtickets verstanden: Flug von A nach B, Umbuchungsmöglichkeiten, Gepäcklimiten, reservierte Plätze etc. (vgl. Kap. 3.1.3). Durch einfache optische (Online-) Anpassungen können Missverständnisse nicht nur bei der Kundensteuerung, sondern bezüglich des *Ticketvalues* verhindert werden (vgl. Kap. 6.5.1).

7.2.2 Kundensteuerung in der Airline-Branche

Airlines steuern das Kundenverhalten über Strafgebühren, Restriktionen oder über das Bonusprogramm (vgl. Kap. 3.1.4). Mit dem Bonusprogramm vermeiden Airlines keine negativen Verhaltensabsichten von Kunden, sondern versuchen bspw., eine höhere Auslastung auf gewissen Strecken zu erreichen. Da vor allem negativ wahrgenommene Ereignisse wie eine Bestrafung ein Aktionssignal aussetzen, gilt es, den Serviceprozess im Hinblick auf diese Situationen zu analysieren (vgl. Abb. 7-1). Innerhalb des Airline-Serviceprozesses wurden aufgrund der qualitativen Empirie die Stufen „Zusatzleistungen vor dem Flug“ und „Check-In / Boarding“ als kritisch bezüglich der Kundenwahrnehmung identifiziert (vgl. Kap. 3.1.2).

Die vorliegenden Ergebnisse der qualitativen Empirie zeigen, dass von Vielfliegern Verspätungen, Überbuchung der Sitzplätze, Kosten für Übergepäck und Umbuchung sowie kompliziertes Check-In als „Bestrafung“ deklariert werden (vgl. Kap. 3.3.4). Die befragten Vielflieger unterscheiden nicht zwischen „kritischen Ereignissen“ oder „mangelhaftem Service“ und „Strafen“. In dieser Arbeit wird jedoch zwischen kritischen Ereignissen und Strafen unterschieden: bei kritischen Ereignissen handelt es sich im Gegensatz zu Strafen um ein Fehlverhalten von Unternehmen (vgl. Kap. 3.4.5).

Die in dieser Arbeit qualitativ untersuchten Airlines, welche das Kundenverhalten verfolgen und dokumentieren, verlassen sich auf das Beschwerdemanagement. Die experimentelle Untersuchung hingegen zeigte, dass Kunden sich nach einer Bestrafung vor allem illoyal verhalten (vgl. Kap. 6.3.7). Illoyales Verhalten wird nicht vom Beschwerdemanagement erfasst. Dieses Ergebnis steht

im Gegensatz zu bestehender Forschung, die negative Word-of-Mouth als ausgeprägtestes Verhalten bei einer Bestrafung postuliert (vgl. Kap. 2.1.4).

7.2.3 Leistungskonformes Verhalten ohne Strafzwang

Straf-(Push-) Massnahmen eignen sich zur Abschreckung unerwünschter Kundensegmente, um das Verhalten von Kunden zu steuern, Kosten zu senken und kurzfristig Umsatz zu erzielen (vgl. Kap. 2.1.2). Zwar sind Push-Massnahmen wirksam zur Kundensteuerung, sie führen aber zu negativen Verhaltensabsichten (vgl. Kap. 6.3.7; 6.5.1). Strafen als Push-Massnahmen werden daher bevorzugt für „kritische“ Stufen im Serviceprozess eingesetzt, wo nicht leistungskonformes Kundenverhalten stark unternehmensschädigend ist (z.B. Übergepäck). Eine Strafandrohung führt zum gewünschten Kundenverhalten und sollte nicht in jedem Fall durch die alternative Pull-Massnahme mit Anreiz-Optionen ersetzt werden (vgl. Kap. 6.4.1).

Anreiz-(Pull-) Massnahmen steuern das Kundenverhalten und ermöglichen gleichzeitig, neue Kundensegmente anzusprechen. Es besteht das Potential zu langfristigem Umsatzwachstum (vgl. Kap. 2.1.2). Pull-Massnahmen sind jedoch schwieriger umzusetzen, da die Wirksamkeit einzelner Massnahmen *ex ante* nicht bekannt ist. Für Dienstleistungsunternehmen besteht die Herausforderung darin, die jeweils geeignete Methode zur Kundensteuerung im Serviceprozess auf verschiedenen Stufen einzusetzen. Konkrete Gestaltungsempfehlungen werden in Kapitel 7.3 erläutert.

Ergebnisse der experimentellen Untersuchung zeigen auf, dass die einzelnen Gebühren (am Check-In) als unfair eingestuft werden (vgl. Kap. 6.3.1). Wird bspw. dieselbe Übergepäckgebühr auf dem Internet als Option angeboten, ist die Nachfrage nach der Option tief. Daraus folgt, dass für die Anreiz-Steuerung über das Internet den Kunden ein Preisnachlass gewährt und kommuniziert werden muss. Im Falle der Übergepäckgebühr liesse sich eine bestimmte Menge an Übergepäck über das Internet zu einem günstigen Preis anbieten (Pull-Massnahme). Dem Kunden würde mitgeteilt, dass dieselbe Menge an Übergepäck am Check-In teurer ist (Push-Massnahme).

Die Kombination von Push- und Pull-Massnahmen ermöglicht die Kundensteuerung über den Preis der Optionen. Sie kann für verschiedene Bereiche der Kundensteuerung eingesetzt werden (Reservierungen, Upgrades, Gültigkeitsdauer von Bonusleistungen im Loyalitätsprogramm etc.). Da der Kunde bereits beim ersten (Online-) Kontakt mit dem Dienstleistungsunternehmen über die Option und die Richtlinie informiert ist, wird er eine Gebühr am Tag des Dienstleistungskonsums nicht als Strafe wahrnehmen. Der Kunde stuft die Massnahme mit einer hohen Fairness-Beurteilung ein, womit der Kundenvorteil erhöht wird. Nicht die kurz-, sondern die langfristige Kundenbeziehung steht im Vordergrund.

Bei nicht leistungskonformem Verhalten sollte die *History* eines Kunden untersucht werden: Was ist der Status des Kunden im Bonusprogramm, was ist sein Kundenwert, hat er sich bereits einmal nicht leistungskonform verhalten? Hierzu müssen neben den bereits bestehenden *Customer Relationship*

Management (CRM-) Daten das Beschwerdeverhalten sowie erfolgte Strafmassnahmen resp. nicht leistungskonformes Verhalten für jeden Kunden aufgeführt werden. Mit diesem Datenmanagement lassen sich Kunden identifizieren, die Einkaufsvereinbarungen mit einer hohen Wahrscheinlichkeit erfüllen. Eine derartige Segmentierung ermöglicht den Einsatz zielgerichteter Kundensteuerungsmassnahmen für spezifische Kundensegmente (z.B. keine Strafen, sondern Anreiz-Optionen). Dabei sind Prinzipien der prozeduralen Gerechtigkeit zu beachten (vgl. Kap. 2.3.2). Um die Mitarbeiter mit Kundenkontakt nicht zu überfordern, empfiehlt sich die Umsetzung der Kundenanalyse mit einem kurzen Hinweis (*Flag*). Das Kunden-*Flagging* kann von verschiedenen Parametern abhängen, die vom Management definiert werden. Die Mitarbeiter wissen bspw. beim Check-In eines Kunden, der ein *Flag* aufweist, dass sie sich bezüglich Bestrafungsmassnahmen gemäss Unternehmensrichtlinien verhalten müssen (vgl. Abb. 7-1).

7.3 Implikationen für das Dienstleistungs- und Airline-Management

Die Forschungsfrage c.1, der wissenschaftliche Beitrag an ein Wirkungsmodell zur Kundensteuerung, wurde mit Kapitel 6.7 beantwortet. Die Forschungsfrage c.2 betrifft die Implikationen für das Dienstleistungs- und Airline-Management zur optimalen Steuerung des Kundenverhaltens (vgl. Kap. 1.4.2): Eine Kombination aus Push- und Pull-Massnahmen, je nach Servicestufe und Situation, führt zur wirksamen Kundensteuerung. Die Massnahmen sollten deutlich kommuniziert, die Wahrnehmung und *History* des Kunden berücksichtigt und ein periodisches *Monitoring* eingeführt werden (vgl. Abb. 7-1).

Mit Kapitel 7.3.1 werden die theoretischen und empirischen Ergebnisse der vorliegenden Arbeit auf das Dienstleistungsmanagement angewendet und ein Gestaltungsmodell zur Umsetzung von Kundensteuerungsmassnahmen erarbeitet. In Kapitel 7.3.2 werden Empfehlungen für das Airline-Management, basierend auf den experimentell untersuchten Szenarien, gemacht. Die Erkenntnisse dieser Arbeit werden mit einer Fallstudie über die fiktive Y-Air praxisnah dargestellt (Kap. 7.3.2).

7.3.1 Dienstleistungsmanagement

Ein faires Kundensteuerungskonzept ist für ein Dienstleistungsunternehmen bei richtiger Umsetzung ein kompetitiver Vorteil und ermöglicht zudem einen höheren, langfristigen Kundenwert. Ziel der Kundensteuerung ist, das Kundenverhalten zu kanalisieren und gleichzeitig die Kundenzufriedenheit nicht zu beeinträchtigen. Basierend auf den zentralen theoretischen und empirischen Ergebnissen aus Kapitel 7.2 werden in Abbildung 7-1 Gestaltungsempfehlungen zur Umsetzung von Kundensteuerungsmassnahmen dargestellt.

Zur Segmentierung von Kunden, die sich nicht leistungskonform verhalten, dient als erster Schritt, neben der *Analyse* der Servicekette und dem CRM, der Kundenwert (vgl. Abb. 7-1). Weist ein Kunde einen hohen Kundenwert aus, sollte das Strafmass angepasst und wenn möglich Anreiz-

Optionen eingesetzt werden. Zur Berechnung des Kundenwerts wird auf Belz und Bieger (2006) verwiesen.

Abbildung 7-1: Gestaltungsempfehlungen zur Umsetzung der Kundensteuerung

	Design der Kundensteuerungsmassnahmen	Instrument
Analyse Ausgangssituation	<ul style="list-style-type: none"> – Kontaktpunkte des Kunden mit dem Unternehmen analysieren. Identifikation nicht leistungskonformen Kundenverhaltens; – Situationen (kritische) an den Kontaktpunkten evaluieren; – Potentiell nicht leistungskonforme Kundensegmente identifizieren. 	<ul style="list-style-type: none"> – Analyse Servicekette – CRM (Beschwerde-management) – Kundenwert
Szenarien erarbeiten	<ul style="list-style-type: none"> – Situationen evaluieren: Branchen-Standard, Kenntnisstand der Kunden, Möglichkeit zur Anreiz-Steuerung, Community-Sicht; – Einteilen der evaluierten Situationen in Situationen, die: <ul style="list-style-type: none"> (a) zu hohen Kosten bei nicht leistungskonformem Verhalten führen; (b) sich während des Kaufprozesses als Optionen gestalten lassen. 	<ul style="list-style-type: none"> – Fokusgruppen – Beobachtung
Umsetzung	<ul style="list-style-type: none"> – Entwicklung geeigneter Push- / Pull-Massnahmen: <ul style="list-style-type: none"> – Für a) Push-, für b) Pull-Massnahmen einsetzen; – Massnahmen je nach Ziel der Unternehmung an die jeweiligen Markt- und Wettbewerbsbedingungen anpassen. – Deutliche Kommunikation der Massnahmen <ul style="list-style-type: none"> – Für Preisgestaltung Equity-Sicht des Kunden berücksichtigen; – Pull-Massnahmen günstiger als Push-Massnahmen offerieren; – Empfehlung: Kombination von Push- und Pull-Massnahmen. – Sicherstellen der Compliance (z.B. mit Flagging) 	<ul style="list-style-type: none"> – Kommunikation über den Internetkanal; – Training von Mitarbeitern und Drittfirmen
Monitoring	<ul style="list-style-type: none"> – Monitoring mit ausgewählten Kunden, wobei das vorliegende Experimental-Setting zur Beurteilung einsetzbar ist. 	<ul style="list-style-type: none"> – Experimentelle Versuchsanordnung

Quelle: Eigene Darstellung

In einem nächsten Schritt werden *Szenarien erarbeitet*, basierend auf der Analyse der Ausgangssituation. Die Szenarien helfen, im Sinne einer Dimensionsanalyse, einen Überblick über negativ wahrgenommene Situationen zu erhalten. Instrumente hierfür sind Fokusgruppen mit Kunden oder eine Beobachtung von Kunden während des Dienstleistungsprozesses. Als Grundlage zur Beobachtung kann das Kategoriensystem aus Kapitel 3.5.3 verwendet werden. Die analysierten Situationen sollten eingeteilt werden in Situationen, die bei nicht leistungskonformem Kundenverhalten zu hohen Kosten führen und bei denen sich Kunden alternativ mit Anreiz-Optionen steuern lassen.

Zur *Umsetzung* der Kundensteuerung wird, falls möglich, eine Kombination von Push- und Pull-Massnahmen empfohlen. Die zunehmende Bedeutung des Online-Buchungsprozesses macht die richtige Kommunikation des Dienstleistungsumfangs und der Optionen sowie Bedingungen unabdingbar. Die Umsetzung von Kundensteuerungsmassnahmen über das Internet erfordert zur Preisgestaltung den Einbezug der Equity-Perspektive. Zudem sollten bei kombiniertem Einsatz von Push- und Pull-Massnahmen die (Online-) Anreiz-Optionen günstiger als Strafgebühren gestaltet werden. Bei der Umsetzung der Kundensteuerungsmassnahmen bleibt zu beachten, dass indirekte Kanäle,

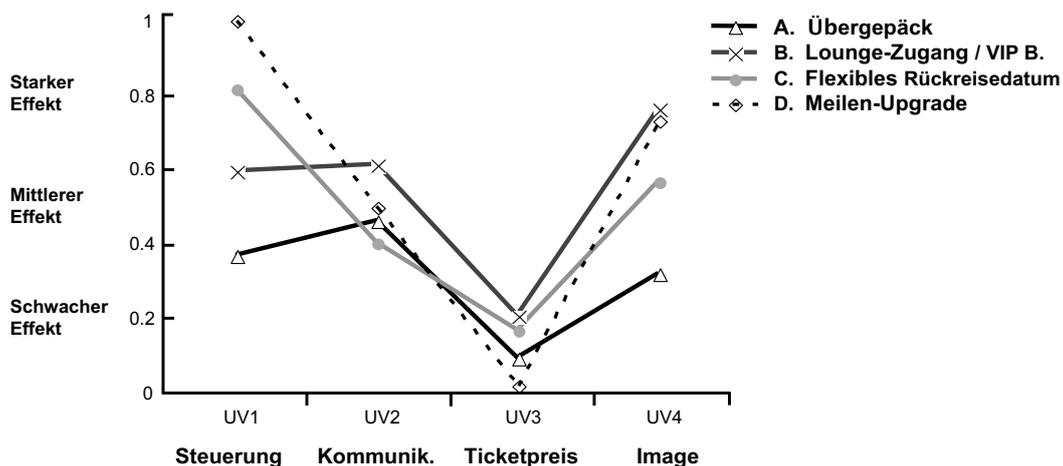
wie bspw. Reisebüros, zur Umsetzung des Kundensteuerungsmanagements einen höheren personellen und finanziellen Aufwand erfordern. Mitarbeiter der Drittfirmen müssen trainiert werden. Der Zugang zu den CRM-Daten sollte für das *Flagging* ermöglicht werden (falls umsetzbar).

Sind die Steuerungsmassnahmen umgesetzt, muss die Wirksamkeit mit einer Fairness-Beurteilung und Abfrage der Verhaltensabsichten periodisch überprüft werden (*Monitoring*). Dabei bietet sich ein Fragebogen mit experimentellem Aufbau an. Als Vorlage kann die in dieser Arbeit eingesetzte Versuchsanordnung eingesetzt werden (vgl. Kap. 5.2).

7.3.2 Airline-Management

Aus der Forschung zu kritischen Ereignissen folgt, dass Flugpassagiere Fehler im Dienstleistungsprozess akzeptieren (vgl. Kap. 3.1.6). Entscheidend ist, wie die Firma auf Fehler reagiert (Wiedergutmachung). Analog zu Forschungserkenntnissen im Bereich der Wiedergutmachung sollte bei einer Strafsituation eine schnelle und adäquate Erläuterung des Strafgrundes erfolgen (vgl. Kap. 3.1.7). Eine späte Reaktion auf Beschwerden von Kunden von Seiten des Dienstleistungsunternehmens impliziert eine ineffiziente Organisation. Zur konfirmatorischen Überprüfung des Forschungsmodells wurden verschiedene Szenarien im Airline-Serviceprozess untersucht (vgl. Kap. 3.8). Die experimentellen Bedingungen als Einflussfaktoren im Modell wurden in jedem Szenario äquivalent manipuliert (vgl. x-Achse in Abb. 7-2).

Abbildung 7-2: Überblick der Effektstärke pro Szenario (Flughafen-Umfrage)



Legende: B. = Boarding, Kommunik. = Kommunikation;
Quelle: Eigene Darstellung

Zur Untersuchung der Einflussfaktoren auf die Fairness-Beurteilung wird in Abbildung 7-2 die Post-hoc-Effektstärke für jedes Szenario einzeln dargestellt (y-Achse). Die Post-hoc-Effektstärke misst die *praktische Bedeutung* verschiedener experimenteller Bedingungen. Sie ist besser geeignet als der Mittelwertsvergleich, weil sie unabhängig von der Stichprobengröße ist. Die Ausprägung des

Effekts ist gemäss der Beurteilung nach Cohen (1988) in „schwacher“ ($d = 0.2$), „mittlerer“ ($d = 0.5$) und „starker“ ($d = 0.8$) Effekt aufgeteilt.

Aus Abbildung 7-2 wird deutlich, dass über alle Szenarien A bis D hinweg ein Trend besteht: Die Kundensteuerung (UV1) hat den stärksten Einfluss auf die Fairness-Beurteilung, gefolgt von der Art der Kommunikation (UV2) sowie dem *ex ante* Image (UV4) der Airline (vgl. Kap. 6.3.1 - 6.3.4). Die hohe Effektstärke bei der Kundensteuerung ist gleichgerichtet, was bedeutet, dass je nach Ausgestaltung der Kundensteuerungsmassnahme ein starker Effekt auf die Fairness-Beurteilung zu erwarten ist. Wird bspw. eine Richtlinie mit Strafe eingesetzt, so hat dies bei einem Szenario mit hoher Effektstärke zur Folge, dass Kunden die Massnahme mit hoher Wahrscheinlichkeit als unfair beurteilen. Wenig Einfluss hat hingegen der Ticketpreis (UV3) als experimentelle Bedingung. Es spielt also keine Rolle, welcher Preis für ein *Economy-Class*-Ticket bezahlt wurde - die Fairness-Beurteilung einer Kundensteuerungsmassnahme bleibt dieselbe. Im Szenario Meilen-Upgrade besteht jedoch ein gegensinniger Interaktionseffekt des Ticketpreises auf die Faktoren Kundensteuerung und Kommunikation (vgl. Kap. 6.4.2).

Aus der vorliegenden experimentellen Untersuchung geht hervor, dass im Szenario „Meilen-Upgrade“ Passagiere besonders ausgeprägte negative Verhaltensaussichten äussern (vgl. Kap. 6.3.7). Zu vermeiden sind Situationen, bei denen dem Kunden keine Wahl gelassen wird und das Unternehmen dem Kunden nicht entgegenkommt (z.B. mit einer nachträglichen Anreiz-Option). Bei der Erläuterung des Strafgrundes muss dem Kunden zudem aufgezeigt werden, dass sein nicht leistungskonformes Verhalten unternehmensschädigend ist.

Die folgenden Empfehlungen gelten für alle Airlines, unabhängig vom eingesetzten Geschäftsmodell. Das Geschäftsmodell als Einflussfaktor auf die Fairness-Beurteilung wurde mit dem Online-Experiment als unwirksam identifiziert (vgl. Kap. 6.5.2).

Übergepäck

Im Szenario Übergepäck wurde untersucht, ob sich die Fairness-Beurteilung zwischen der Online- und der Offline-Gruppe unterscheidet (vgl. Kap. 4.3.1). Analysiert wurde die Kundenwahrnehmung, je nachdem, ob fünf Kilo Übergepäck am Check-In bezahlt werden musste oder als Option bei der Online-Buchung bezahlt werden konnte (CHF 150.-). Da die Preise für Treibstoff und Frachtgut eine steigende Tendenz aufweisen, wird die Bedeutung eines optimal beladenen Flugzeugs grösser. Das Gewicht des Passagiergepäcks spielt dabei eine wichtige Rolle.

Die Übergepäckgebühr ist die Kundensteuerungsmassnahme, deren praktische Bedeutung bezüglich Fairness-Rating am schwächsten ausgeprägt ist (vgl. Abb. 7-2). Mit einem mittelstarken Effekt ist das Differenzmass nach Cohen (1988) für ein neues Forschungsgebiet ausreichend.

Die detaillierte Auswertung zum Antwortverhalten bei Übergepäck zeigt, dass am Check-In die Zahlungsbereitschaft im Vergleich zum Online-Szenario höher ist (42% Check-In vs. 10% Online-Nachfrage). Die hohe Ablehnungsrate im Online-Übergepäck-„Angebot“ liegt gemäss Probanden-

antworten vor allem am zu teuren Preis für fünf Kilo Übergepäck (CHF 150.-). Die Untersuchung der Effekte zeigt aber auf, dass eine mittelstarke, gleichsinnige Korrelation mit der Kommunikation besteht. Als Implikation folgt, dass Übergepäck effizient über Richtlinien zu steuern ist, wobei eine deutliche (Online-) Kommunikation erfolgen muss, um die Gerechtigkeitswahrnehmung nicht negativ zu beeinflussen.

Die signifikanteste Verhaltensreaktion im Szenario Übergepäck ist im Gegensatz zu den anderen Szenarien die „Beschwerde“, wobei insbesondere *Gold Member* im Vielfliegerprogramm dazu neigen. Eine Segmentierung der Übergepäck-Bestrafungsrichtlinien nach Status im Vielfliegerprogramm drängt sich auf. Die übrigen flugdemografischen Variablen hatten keinen Einfluss auf die Fairness oder Akzeptanz einer Übergepäckgebühr.

Lounge-Zugang und VIP Boarding

Die Probanden wurden in diesem Szenario vor die Wahl gestellt, für CHF 40.- einen Lounge-Zugang oder Boarding mit Priorität online zu erwerben oder nicht. Im Falle von „Bestrafung“ wurde den Probanden mitgeteilt, dass sie diese Anrechte am Check-In nicht haben, aber im separaten Airline-Office für CHF 40.- kaufen können.

Der Zugang zur Lounge und das Boarding mit VIP-Status eignen sich als Massnahmen zur Anreizsteuerung über das Internet. Am Check-In war die Nachfrage nach dieser Option (CHF 40.-) im Gegensatz zum Online-Szenario tiefer (12% Check-In vs. 19% Online-Nachfrage).

Eine starke Bedeutung für das Fairness-Rating eines Lounge-Zugangs und VIP Boarding hat das Image der Airline. Daraus folgt, dass für Passagiere das Konstrukt Airline-Image u.a. über den Lounge-Zugang und VIP Boarding definiert wird.

Fehlt der Lounge-Zugang oder das VIP Boarding, akzeptieren vor allem *Economy-Class*-Passagiere diesen Umstand nicht und verhalten sich illoyal. Interpretiert wird diese Tatsache mit den Wahlmöglichkeiten, die *Business-Class*-Passagiere haben. Die meisten *Business-Class*-Passagiere waren in der vorliegenden Untersuchung Mitglieder bei mehreren Vielfliegerprogrammen. Sie wissen, dass sie bei fehlendem Lounge-Zugang eine Ausweichmöglichkeit haben (vgl. repräsentative Antworten in A.5.3). Wenn ein *Business-Class*-Passagier negatives Verhalten äusserte, war dies die Beschwerde. Bezüglich Flugerfahrung zeigte sich bei geschäftlich Reisenden, dass sie sich bei geringer Flugerfahrung illoyal verhalten. Passagiere mit viel Flugerfahrung wissen, dass Fehler im Dienstleistungsprozess auftreten können, und verfügen über Erfahrung im Umgang mit diesen Situationen.

Flexibles Rückreiseticket

Ähnlich wie im Szenario Lounge-Zugang konnten die Probanden der Online-Gruppe ein flexibles Rückreiseticket online für CHF 200.- erwerben. Es wurde experimentell untersucht, wie sich die Wahrnehmung und das Verhalten unterscheiden, wenn den Probanden am Check-In erklärt wird,

dass ihr Ticket keine flexible Wahl des Rückreisedatums ermöglicht. Den Probanden der Offline-Gruppe wurde aber die Wahl gelassen, für CHF 200.- das flexible Rückreiseticket am separaten Airline-Schalter zu kaufen.

Passagiere bezahlen zwar das flexible Rückreiseticket am Check-In, aber sie empfinden die Gebühr als unfair. Das konkrete Verhalten beim Kauf eines flexiblen Rückreisetickets für CHF 200.- ist bezüglich der Kundensteuerung unklar. 36% der Probanden zahlen die Gebühr am Check-In-Schalter, 40% bezahlen sie nicht. Im Online-Szenario interessieren sich nur 14% für diese Option, wobei für 26% der Preis zu teuer ist. Wird hingegen auf die Fairness-Beurteilung geachtet, kann mit der Steuerung über das Internet eine hohe Fairness-Wahrnehmung erreicht werden.

Je höher die Flugerfahrung von Privatreisenden, umso höher ist die Akzeptanz bezüglich einer Richtlinie zum Rückreisedatum. Geschäftsreisende weisen aufgrund der benötigten Flexibilität ein tiefes Akzeptanzniveau auf.

Meilen-Upgrade

Als „strengstes“ Szenario wurde den Probanden der Offline-Gruppe erklärt, dass der beabsichtigte Upgrade in die nächsthöhere Buchungsklasse mit Meilen aus organisatorischen Gründen nicht möglich ist. Im Gegensatz zu den übrigen Szenarien bestand keine „Option“ am Check-In. Ebenfalls untersucht wurde, wie stark ein Upgrade in die nächsthöhere Klasse während des Online-Buchungsprozesses nachgefragt wird.

Das Meilenprogramm eignet sich zur Kundensteuerung, da die praktische Bedeutung (Effekt) verschiedener Kundensteuerungsmassnahmen ausgeprägt ist (vgl. Abb. 7-2). Zudem zeigt die Auswirkung auf die Variable *ex post* Image, wie sensitiv Probanden auf Bonusprogramme reagieren (resp. auf ein nicht gewährtes Meilen-Upgrade).

Da im Szenario Meilen-Upgrade als einziges Szenario keine Option am Schalter angeboten wurde, sind die kognitiven und Verhaltensreaktionen am stärksten ausgeprägt. Der hohe Effekt beim *ex ante* Image bedeutet nicht nur einen positiven Einfluss eines guten Images auf die Fairness-Beurteilung. Er zeigt auch auf, dass eine Airline mit einer tieferen Image-Wahrnehmung bei der Implementierung von Kundensteuerungsmassnahmen vorsichtig vorgehen muss. Ansonsten wird die negative Image-Wahrnehmung verstärkt. Dies gilt ebenfalls im Szenario Lounge-Zugang und VIP Boarding.

Der nachgewiesene Interaktionseffekt zwischen dem Ticketpreis und der Kommunikation sowie der Kundensteuerung bedeutet, dass eine deutlich kommunizierte Anreiz-Steuerung, kombiniert mit einem tiefen Ticketpreis, zu einem hohen Fairness-Rating führt (vgl. Kap. 6.4.2). Wird hingegen eine Richtlinie unklar kommuniziert, wirkt der Ticketpreis im Szenario „Meilen-Upgrade“ gemäss Interaktionsdiagramm gleichläufig: das Fairness-Rating nimmt mit der Höhe des Ticketpreises zu. Die Ursache dieses auf den ersten Blick unklaren Zusammenhangs kann mit den Werbemassnahmen von Airlines erklärt werden. Konsumenten wurden vielfach von Airlines mit tiefen Preisen angelockt, wobei diese Preise keine Flughafentaxen und sonstige Gebühren beinhalteten (o.V.,

2007b). Als im Experiment sehr tiefe, unklar kommunizierte Ticketpreise „angeboten“ wurden, fiel das Fairness-Rating aufgrund der Probanden-Erfahrungen mit Airline-Werbemassnahmen tief aus. Die Probanden assoziierten intransparent kommunizierte Preise mit unfairen Airline-Praktiken. Beim Meilen-Upgrade sind es wiederum die *Economy-Class*-Passagiere, die einen nicht gewährten Meilen-Upgrade am Check-In weniger akzeptieren als *Business-Class*-Passagiere. Ein nicht gewährter Meilen-Upgrade ist imageschädigend.

Folgerung

Gerade in der Airline-Branche, wo Bestrafungsrichtlinien zum Branchenstandard gehören, kann sich eine Airline mit einem ganzheitlichen Kundensteuerungskonzept differenzieren. Auf vielen Strecken herrscht wenig Konkurrenz, weshalb Kunden Bestrafungen hinnehmen. Nehmen jedoch Konkurrenzangebote zu, neigen diejenigen Passagiere, die sich nicht fair behandelt fühlen, zum Anbieterwechsel.

Aus theoretischen Erkenntnissen zum Bonusprogramm folgt, dass der Status und die Exklusivität die treibenden Faktoren für die Teilnahme am Programm sind (vgl. Kap. 2.1.1). Airline-Passagiere haben aufgrund der verschiedenen Beförderungsklassen und Bonusprogramme gelernt, in Statuskategorien zu denken (vgl. Kap. 3.1.4). Diese Eigenschaften lassen sich für die Kundensteuerung einsetzen. Sind Kundensegmente definiert, die sich mit hoher Wahrscheinlichkeit leistungskonform verhalten, können sie mit Anreiz-Optionen, die einen gewissen Status oder Exklusivität versprechen, gesteuert werden. Kundensegmente, die sich mit hoher Wahrscheinlichkeit nicht leistungskonform verhalten, werden bevorzugt über Bestrafungsrichtlinien gesteuert.

Gemäss der *Prospect Theory* sehen Kunden Einkaufsentscheide als Entscheide zwischen Gewinnen und Verlusten. Bestrafungsfairness kann demnach durch die Kommunikation der Massnahmen beeinflusst werden (McCarthy und Fram, 2000). Je nachdem, wie eine Strafe kommuniziert oder *geframed* wird, verhält sich der Durchschnittskunde anders. Im Airline-Bereich eignet sich das bereits bestehende Bonusprogramm zur Umsetzung der Kundensteuerung. Die Ergebnisse der vorliegenden experimentellen Untersuchung bestätigen die von der Forschung postulierte hohe Sensitivität von Kunden bezüglich Bonusprogrammen und deren Incentivierungs- sowie Einlösemechanik. Es liessen sich bspw. folgende zwei äquivalente Szenarien empirisch von Flugreisenden auf ihre Präferenz hin bewerten (vgl. Tab. 7-1). In Szenario „Strafe“ wird gemäss *Prospect Theory* die Strafe von 15'000 Meilen als Verlust angesehen, wogegen in Szenario „Anreiz“ ein potentieller Gewinn von 15'000 Meilen besteht. Obwohl beide Szenarien bezüglich des „Totalpreises“ äquivalent sind, bevorzugen Individuen Szenario „Anreiz“. Voraussetzung für diese Art der Kundensteuerung ist, dass der Kunde Mitglied im Bonusprogramm ist und über genügend Meilen verfügt. Zudem darf der Preis des Gratisflugs im Vergleich zu konkurrierenden Vielfliegerprogrammen nicht zu teuer werden. Dieses *Framing* der Kundensteuerung lässt sich auch ohne Meilenprogramm, über reale Strafgebühren und Anreiz-Optionen implementieren (vgl. Fallstudie Y-Air).

Tabelle 7-1: Framing der Kundensteuerung bei einer Flugreise

	Kundensteuerung	Bedingungen
Szenario „Strafe“	Ein Europaflug kostet 30'000 Meilen. Wenn der Flug abgesagt oder geändert wird, muss eine Strafe von 15'000 Meilen bezahlt werden.	- Preis: 30'000 Meilen - Strafe: 15'000 Meilen - Total bei Antritt Reise: 30'000 Meilen
Szenario „Anreiz“	Ein Europaflug kostet 45'000 Meilen. Wenn der Flugreisende den Flug gemäss originärem Flugschein durchführt, werden ihm 15'000 Meilen rückerstattet. Ansonsten kann der Flug ohne Zusatzgebühr abgesagt oder geändert werden.	- Preis: 45'000 Meilen - Anreiz: 15'000 Meilen - Total bei Antritt Reise: 30'000 Meilen

Quelle: Eigene Darstellung

Unternehmen sollten gemäss der *Prospect Theory* mehrere „Strafen“ in Kombination kommunizieren (nicht in verschiedenen Schritten). Wenn bspw. ein Mitglied des Vielfliegerprogramms seinen Status nicht halten kann und gleichzeitig die Gültigkeit der noch bestehenden Meilen verfällt, sollten beide nicht monetären Bestrafungen gleichzeitig kommuniziert werden. Die Fluggesellschaft könnte auch das *Downgrading* des Vielflieger-Status (Bestrafung) mit einem Meilenbonus (Gewinn) kombinieren und somit das *Downgrading* möglichst angenehm gestalten. Es ist auch denkbar, dass eine Strafzahlung durch die Airline erlassen wird, wenn der Kunde sich verpflichtet, eine bestimmte Anzahl Flüge während einer gewissen Zeit durchzuführen (ähnlich den „Statusmeilen“ bei Vielfliegerprogrammen). Schwieriger umzusetzen sind hingegen der Wegfall einer Bestrafung bei unvorhersehbaren Ereignissen von Kunden wie z.B. Krankheit. Diese Ereignisse werden bereits heute über Versicherungen abgedeckt.

Über alle Szenarien hinweg betrachtet zeigt sich, dass Passagiere in Situationen, die wenig Wahlmöglichkeit lassen, über Strafgebühren gesteuert werden können. Die Übergepäckgebühr und das flexible Ticket wurden als *ex ante* Option wenig nachgefragt, während die Probanden am Check-In zwar eine tiefe Fairness-Beurteilung zeigten, aber diese Optionen *ex post* schliesslich bezahlten. Die Kundensteuerung hängt stark von der Ausgestaltung des Preises der Optionen ab, wobei Anreiz-Optionen über das Internet günstiger sein müssen als dieselben Optionen am Check-In. Das *ex ante* Image einer Airline wirkt bei der Umsetzung der Kundensteuerungsmassnahmen unterstützend, falls die Airline als Qualitätsunternehmen wahrgenommen wird. Eine Segmentierung der Kundensteuerungsmassnahmen nach Produktgruppen (*Economy*, *Business* und *First Class*) oder Flugerschaft sollte je nach Möglichkeit, Szenario und Reiseart (geschäftlich, privat) vorgenommen werden. Über alle Szenarien hinweg betrachtet konnte kein systematischer Einfluss der demografischen Variablen auf die Akzeptanz und die Verhaltensabsichten nachgewiesen werden.

Da viele Kontaktpunkte im Airline-Serviceprozess nicht direkt von der Airline durchgeführt werden, sind die intensive Schulung der Mitarbeiter und Qualitätssicherung von Drittfirmen nötig (z.B.

Abfertigungsfirmen), falls die Kundensteuerung nicht nur über den Online-Buchungsprozess eingeführt werden soll.

Entwurf eines fiktiven Kundensteuerungskonzepts für Y-Air



Schritt 1: Analyse Ausgangssituation

Y-Air positioniert sich als Qualitätsairline mit Fokus auf Langstreckenflüge. In letzter Zeit hat sich die Konkurrenzsituation für Y-Air auf einzelnen Strecken verschärft. Zusätzlich schneidet Y-Air bei Reputationsmessungen von Kunden im Vergleich zum Vorjahr schlecht ab. Dies widerspiegelt sich in zunehmenden Kunden-Beschwerden und damit verbundenen Kompensationskosten. Das Management vermutet, dass sich Kunden illoyal verhalten und zur Konkurrenz wechseln. Es möchte mit der Implementierung eines Kundensteuerungskonzepts die Kundenzufriedenheit erhöhen und gleichzeitig die Kosten senken.

Schritt 2: Szenarien erarbeiten und Ziel definieren

Die Beschwerdeanalyse innerhalb der Servicekette zeigt, dass viele Kunden mit dem Internetauftritt und dem Buchungsprozess von Y-Air unzufrieden sind. Das Online-Buchungsverhalten wird mit einer Fokusgruppe und anhand der Abbruchraten während des Buchungsprozesses analysiert. Bemängelt werden die Struktur der Internetseite und die unklare Kommunikation der Tarifbedingungen. Die meisten Probanden stufen den Buchungsprozess auf einer Rating-Skala als „unfair“ ein.

Die Folgen der ungenügend kommunizierten Tarifbedingungen auf dem Internet treten am Check-In auf. Insbesondere Übergepäck führt aufgrund hoher Strafgebühren (Push-Massnahmen) zu unzufriedenen Passagieren am Check-In. Gespräche mit Mitarbeitern der Abfertigungsfirmen zeigen, dass Kunden am Check-In bezüglich der Y-Air-Übergepäck-Limiten und -Strafgebühren unvollständig informiert sind. Zudem ist der Y-Air-Umsatz im lukrativen Frachtgeschäft gesunken, da auf Linienflügen aufgrund von Passagier-Übergepäck die Zuladung zusätzlicher Fracht eingeschränkt ist.

Ein weiteres Ergebnis der kundenseitigen Beschwerdeanalyse und der Fokusgruppen ist die hohe Unzufriedenheit von Vielfliegern mit den Kosten für die Umbuchung eines Fluges und den Möglichkeiten zum Einlösen von Y-Air-Meilen im Loyalitätsprogramm. Die Y-Air-Vielflieger, die zunehmend und kurzfristig Flüge umbuchen, verursachen dem Unternehmen operationelle Kosten.

Ziel des Y-Air-Kundensteuerungssystems ist, Kunden für leistungskonformes Verhalten zu motivieren, um Kosten einzusparen. Gleichzeitig soll die Y-Air-Image-Wahrnehmung bei Kunden erhöht werden. Hierzu soll für das als kritisch definierte Szenario „Übergepäck“ eine Kombination von Push- und Pull-Massnahmen eingesetzt werden. Mit der Implementierung von Anreiz-Optionen (Pull-Massnahmen) im Szenario „Loyalitätsprogramm“ sollen neue Möglichkeiten im Loyalitätsprogramm zur Incentivierung und Einlösemechanik von Meilen geschaffen werden sowie die flexible Umbuchung erlaubt sein.

Schritt 3: Umsetzung der Kundensteuerung

Y-Air will sich im Vergleich zur Konkurrenz mit einem umfassenden Kundensteuerungskonzept als passagierfreundliche Airline positionieren. Hierzu werden die neuen Kundensteuerungsmassnahmen mit verschiedenen Werbemassnahmen bei der Zielgruppe vorgestellt.

Szenario „Übergepäck“: Das ermittelte Szenario wird neu über den Online-Buchungsprozess umgesetzt. Internet-Benutzer können die Übergepäck-Option (5 oder 10 kg) bei der Preisübersicht auswählen. Diese Pull-Massnahme wird klar und deutlich auf der Internetseite kommuniziert. Gleichzeitig werden die Internet-Benutzer darauf aufmerksam gemacht, dass dieselbe Option am Check-In 30% teurer ist (Push-Massnahme). Y-Air hat erkannt, dass das Verhältnis zwischen dem Ticketpreis und dem Preis der Pull-Massnahme auf dem Internet, resp. der Push-Massnahme am Check-In, für die Kundenzufriedenheit entscheidend ist. Basierend auf einer experimentellen Untersuchung am Heimflughafen von Y-Air wird das optimale Verhältnis zwischen dem Anreiz-Preis / der Strafgebühr und dem Ticketpreis festgelegt. Dabei werden die Steuerungsziele von Y-Air berücksichtigt.

Szenario „Loyalitätsprogramm“: Aufgrund theoretischer Erkenntnisse wird angenommen, dass mit Anreiz- (Pull-) Massnahmen weniger Umbuchungen erfolgen und damit Kosten gesenkt werden können. Um die Unzufriedenheit bei den Vielfliegern bezüglich der Umbuchung zu verringern und gleichzeitig die Attraktivität des Y-Air-Loyalitätsprogramms zu erhöhen, werden eine neue Preispolitik und ein neues Einlösesystem eingeführt. Neu ist im Flugpreis (z.B. 80'000 Y-Air-Meilen oder CHF 800.-) das Recht inbegriffen, den Flug einmal umzubuchen. Macht der Passagier von dieser Online-Option keinen Gebrauch, werden ihm 10'000 Meilen elektronisch auf sein Meilenkonto zurückerstattet oder CHF 100.- für den nächsten Flug gutgeschrieben.

Y-Air hat sich zum Ziel gesetzt, weitere Anreiz-Optionen einzusetzen. Als umsetzbare Anreiz-Optionen über den Internet-Buchungsprozess werden identifiziert:

- Anzahl Gepäckstücke oder Sperrgüter (z.B. Fahrrad, Surfbrett)
- Lounge-Zugang und / oder VIP Boarding
- Sitzplatzreservation und / oder spezielle Verpflegung
- Upgrade in die nächsthöhere Klasse

Schritt 4: Monitoring

Das Monitoring der Kundensteuerungsmassnahmen beinhaltet die Analyse des Nachfrageverhaltens über das Internet, vor und nach Einführung der Anreiz-Optionen, sowie das Kunden-Flagging. Das Kunden-Flagging wird auf der Servicestufe Check-In eingeführt. Es werden verschiedene Status-Flags definiert, die von den Parametern „Anzahl Beschwerden“, „Anzahl geleisteter Strafgebühren“ und „Status im Vielfliegerprogramm“ abhängen. Die Mitarbeiter am Check-In werden geschult und lernen, das Y-Air-Kunden-Flagging im CRM zu kontaktieren. Wichtige und kritische Kundensegmente erhalten bspw. ein „Red Flag“. Die Check-In-Mitarbeiter werden angewiesen, diese Kundensegmente bei Zwischenfällen an den Supervisor weiterzuleiten. Sofern Möglichkeiten bestehen, wird der Supervisor gemäss der Y-Air-Kundensteuerungs-Checkliste kulant auf die Anfrage eines Red-Flag-Passagiers reagieren. Passagiere werden darauf hingewiesen, dass sie sich nicht leistungskonform verhalten haben.

7.4 Kritische Würdigung der Untersuchungsergebnisse

7.4.1 Forschungsmodell

Das Image als multidimensionales Konstrukt wurde im vorliegenden Forschungsmodell nur rudimentär abgebildet und gemessen, da es nicht im Zentrum der Forschungsfrage stand. Um das *ex ante* und *ex post* Image im Modell genauer zu erfassen, wird eine vertiefte Analyse des Konstrukts Image im Modell vorgeschlagen. Insbesondere weil eine Auswirkung auf die Image-Wahrnehmung aufgrund der Kundensteuerungsmassnahmen nachgewiesen werden konnte.

Das Konstrukt Fairness-Beurteilung wurde als zentrales Konstrukt zur Messung von Strafmassnahmen identifiziert. Entgegen aktuellen Forschungsergebnissen im Dienstleistungsmarketing konnte die ursächliche Wirkung emotiver Reaktionen auf die Verhaltensabsichten nicht bestätigt werden. Die kognitive Fairness-Beurteilung kann einen Sachverhalt genauso vorhersagen.

Mit den insgesamt dreizehn getesteten (Haupt-) Hypothesen wurden 40 Szenarien einzeln überprüft. In 38 Szenarien waren die Ergebnisse pro Hypothese äquivalent. Diese hohe Übereinstimmung zeigt, dass das überprüfte Forschungsmodell auf verschiedene Dienstleistungskontexte angewendet werden kann.

Die Preise der Szenarien A bis D für Strafen und Anreize wurden qualitativ erarbeitet und überprüft. Dennoch könnte ein systematischer Einfluss eines zu hohen oder zu tiefen Preises einzelne Implikationen der vorliegenden Arbeit verfälscht haben. Zusätzliche Forschung mit verschiedenen Preisen für Strafen oder Anreize als experimentelle Bedingungen könnte Klarheit schaffen.

7.4.2 Methoden-Kombination

Es wurden zwei verschiedene Datenerhebungsmethoden zur quantitativen Untersuchung des Forschungsmodells eingesetzt. Wie in Arbeiten zur experimentellen Forschung postuliert, war die Abbruchrate bei der Online-Untersuchung höher. Trotzdem widerspiegeln die Ergebnisse der Online-Umfrage die Erkenntnisse aus der schriftlichen Flughafen-Umfrage. Auffallen ist, dass die Effekte der experimentellen Bedingungen im Online-Experiment stärker ausgeprägt waren als die Effekte im Flughafen-Experiment. Für weitere Forschungsarbeiten kann somit auf die aufwendige Untersuchung vor Ort verzichtet werden, solange der Fragebogen resp. das zu untersuchende Forschungsmodell nicht zu umfangreich ist.

Aufgrund der Analyse der demografischen Daten hat das Internet-Experiment nicht die erwartete Ausweitung der Untersuchung auf verschiedene Kulturkreise ermöglicht. Das Flughafen-Experiment war diesbezüglich heterogener und erlaubte neben den Vorteilen einer dem Feldexperiment nahen Untersuchung eine gute Stichprobenauswahl. Zudem hat sich gezeigt, dass Passagiere beim Warten an den Gates für das Ausfüllen des Fragebogens motiviert werden können. Die Störquellen wie Lärm oder Ablenkung durch Drittpersonen sind minim.

Die Befunde der vorliegenden Arbeit werden nicht dadurch entwertet, dass die Experimente letztlich etwas bestätigen, was man vom Alltag her zu wissen meint. Zum einen ist das vermeintliche Alltagswissen nicht immer richtig, und zum anderen kann ein solcher Untersuchungsansatz sehr viel genauere Auskünfte über das Ausmass und Effekte geben, als dies ohne die Durchführung einer solchen Untersuchung möglich wäre.

7.5 Weiterer Forschungsbedarf

Mit der vorliegenden Arbeit wurde das Konstrukt Strafe als Push-Massnahme gegenüber Pull-Massnahmen abgegrenzt. Somit wurde eine Alternative zu den etablierten Kundenbestrafungen erarbeitet und wurden Strategien zu adäquatem Verhalten bei negativem Kundenverhalten empfohlen. Die erstmalige Anwendung auf die Airline-Branche ermöglichte die Ableitung von Implikationen für das Dienstleistungs- und Airline-Management. Als praktische Implikation konnten in vier verschiedenen Szenarien im Airline-Dienstleistungsprozess Handlungsanweisungen gegeben werden. Der Distributionskanal Internet wurde als effiziente Servicestufe für das Kundensteuerungsmanagement identifiziert. Das aus theoretischen Erkenntnissen erarbeitete Wahrnehmungs- und Verhaltensmodell hat sich zur Untersuchung verschiedener exogener und endogener Einflüsse auf die Fairness-Beurteilung bewährt. Die Forschungsfragen a.1 bis c.2 aus Kapitel 1.4.2 wurden beantwortet und das Forschungsziel wurde erreicht. Für weitere Forschung im Bereich von Kundensteuerungsmassnahmen zeichnen sich folgende Ansatzpunkte ab:

- *Untersuchung des vorliegenden Modells anhand einer anderen Branche:* Bereits erwähnt wurden Bestrafungssituationen im Banken-Sektor. Das vorliegende Wirkungsmodell liesse sich anhand neu erarbeiteter Szenarien im Bereich von Bankdienstleistungen oder anderen Transportunternehmen untersuchen. Die Ausweitung einer Untersuchung auf die Konsumgüterbranche würde zur Abgrenzung und Klassifizierung des Konstrukts Bestrafung beitragen.
- *Einbindung neuer experimenteller Bedingungen im Forschungsmodell:* Im Modell wurden aus forschungsökonomischen Gründen nicht mehr als fünf unabhängige Variablen untersucht. Die Wahrnehmung von Kunden wird aber nicht ausschliesslich von diesen Faktoren beeinflusst, sondern von weiteren, kontext-abhängigen Bedingungen wie bspw. der Anwesenheit von Drittpersonen, Stresslevel der Mitarbeiter und Kunden, Stufe im Serviceprozess etc. Zudem kann das Konstrukt Image, wie in Kapitel 7.4.1 erläutert, mehrdimensional abgebildet werden.
- *Vertiefende Analyse der Szenarien und preisliche Anpassung der Strafgebühren und Anreiz-Optionen:* Mit den vier vorgeschlagenen Szenarien zur experimentellen Untersuchung des Forschungsmodells wurde eine Auswahl getroffen. Interessant wäre eine weitere experimentelle Untersuchung zur Aufdeckung von Analogien zwischen verschiedenen Szenarien. Bei-

spielsweise liesse sich untersuchen, ob der Preis einer Strafe oder Option einen systematischen Einfluss auf die Verhaltensabsichten hat. Zudem könnten weitere Service-Stufen im Airline-Bereich untersucht werden, wie z.B. die Stufe *On-Board* (Flugreise).

- *Integration weiterer Distributionskanäle:* Zur Kommunikation von Kundensteuerungsmassnahmen wurde lediglich der Distributionskanal Internet untersucht. Weitere Forschungsfelder eröffnen sich bspw. beim indirekten Vertrieb über Reisebüros oder Call Center. Es liessen sich weitere Implikationen zum Management von Kundensteuerungsmassnahmen ableiten.
- *Longitudinale Studien mit realen E-Commerce-Daten:* Einzelne Airlines haben begonnen, einzelne Anreiz-Optionen über das Internet anzubieten. Daraus ergeben sich interessante Forschungsfelder, falls Zugriff auf diese Daten möglich ist. Am realen Forschungsobjekt könnte bspw. untersucht werden, ob sich das Kundenverhalten nach Einführung neuer Anreiz- oder Strafmassnahmen geändert hat. Zudem liesse sich der Einfluss der Kommunikation (inhaltlich und optisch) untersuchen: Wie muss eine Strafe oder ein Anreiz formuliert sein, damit leistungskonformes Verhalten erreicht wird.
- *Interaktionseffekt des Ticketpreises:* Mit dem Flughafen-Experiment wurde der Interaktionseffekt des Ticketpreises auf die Faktoren Kundensteuerung und Kommunikation untersucht. Nachgewiesen wurde ein disordinaler Interaktionseffekt im Szenario Meilen-Upgrade. Da in den restlichen Szenarien keine Interaktion nachgewiesen werden konnte, drängt sich weiterer Forschungsbedarf auf. Zudem konnte die Hypothese, dass der Ticketpreis einen signifikanten Einfluss auf die Fairness-Beurteilung hat, nur mit einer tiefen Teststärke abgelehnt werden.
- *Kosten-Nutzen-Rechnung:* Eine quantitative Untersuchung einzelner Szenarien könnte die Wirksamkeit der Kundensteuerungsmassnahmen beziffern. Es liessen sich bspw. die Kosten der Umsetzung von Übergepäckstrafen / -optionen mit ihrem Nutzen vergleichen.

A Anhang

A.1 Theoretisch-konzeptionelle Grundlagen

Funktion der Strafe in der Schweiz, Deutschland und den USA

Für das Schweizer Bundesgericht hat die Strafe in erster Linie nicht der Vergeltung, sondern im Rahmen des Verschuldens der Verbrechensverhütung zu dienen, wobei den spezialpräventiven Bedürfnissen insoweit der Vorrang eingeräumt werden soll, wie die generalpräventiven Mindestanforderungen noch gewahrt sind (Kunz, 2003). In Deutschland, wo bis etwa 1965 die Vergeltungstheorie vorherrschte, hat sich inzwischen das Resozialisierungsstrafrecht durchgesetzt (Stratenwerth, 2005). Symptome dafür sind die Ausweitung des bedingten Strafvollzugs und die Einführung der sozialtherapeutischen Anstalt für Täter mit Persönlichkeitsstörungen. In den USA hingegen ist nach anfänglichem *Zweckstrafrecht*, wo bspw. Strafen und psychiatrische Massnahmen angewendet wurden, in letzter Zeit eine Rückkehr zum Vergeltungsstrafrecht zu beobachten (Trechsel und Noll, 2004). Dies wird auch in der amerikanischen Literatur zum Thema Bestrafung erwähnt. Dort wird Vergeltung im Gegensatz zur Abschreckung oder Wiedereingliederung als das wirkungsvollere Konzept beschrieben, weil Vergeltung versucht, ein soziales Gleichgewicht wiederherzustellen. Demgegenüber nützt Abschreckung nur, um dieselbe kriminelle Handlung in Zukunft zu vermeiden (Hart, 1982).

Eigenschaften von Dienstleistungen

Intangibilität ist eine der wichtigsten Eigenschaften von Dienstleistungen: sie haben keine physische Dimension. Die Dienstleistung einer Airline ist nicht das physische Flugzeug, sondern die intangible Flugreise (Stafford und Day, 1995; Shostack, 1977). Die meisten Dienstleistungen benutzen tangible Güter, um eine Leistung bereitstellen zu können (Berry, 1980). Mit *Inseparabilität* ist der Zusammenfall von Produktion und Konsum bei einer Dienstleistung gemeint (*Uno-Actu-Prinzip*). Güter werden zuerst produziert und anschliessend konsumiert (Bieger, 2003; Meffert und Bruhn, 2000). Weil der Leistungsprozess bei Dienstleistungen menschliche Interaktion voraussetzt und sich diese von Interaktion zu Interaktion unterscheidet, entsteht *Heterogenität* (Gabbott und Hogg, 1994; Abernethy und Butler, 1992). Dienstleistungen werden von menschlicher Interaktion bestimmt, weshalb jedes Dienstleistungsprodukt anders aussieht. *Vergänglichkeit* hingegen bedeutet, dass Dienstleistungen nicht gelagert werden können. Sie haben eine Echtzeit-Komponente (Lovelock, 1992). Wenn sich also die Nachfrage nach einer Destination kurzfristig ändert, kann bspw. nicht auf ein Lager von Flugreisen zurückgegriffen werden. Mit dem Verkauf eines Gutes geht das *Eigentum* in den Besitz des Käufers über. Bei Dienstleistungen aber wird von Konsumenten ein Anrecht auf Gebrauch gekauft (Kotler, 2006). Im Beispiel einer Flugreise hat der Kunde ein Anrecht auf den Nutzen der Flugreise, aber nicht auf die Dienstleistung selbst. Das Fehlen von Eigentum betont die endliche Natur von Dienstleistungen (Gabbott und Hogg, 1994).

A.2 Grundzüge der Empirie

A.2.1 Beobachtung

Die Reliabilität einer Beobachtung als Messinstrument kann bestimmt werden, indem verschiedene Beobachtungen miteinander verglichen werden. Eine Messung wird mehrfach unter theoretisch äquivalenten Bedingungen wiederholt. Wenn die Unterschiede zwischen den verschiedenen Messungen gross sind, ist das Messinstrument fehleranfällig (Greve und Wentura, 1997). Eine hohe Reliabilität ist die Voraussetzung der Validität: Wird tatsächlich das gemessen, was gemessen werden sollte? Greve und Wentura (1997) weisen in diesem Zusammenhang darauf hin, dass Validitätskriterien häufig verletzt werden, weil auf bestehende Sammlungen von Kategoriensystemen / Verfahren zur Beobachtung zurückgegriffen wird. Die Validität kann überprüft werden, wenn im Sinne einer Triangulation die Ergebnisse der Beobachtung mit Ergebnissen aus anderen Untersuchungen verglichen werden.

Beobachter und das Kategoriensystem bilden zusammen ein Messinstrument, das möglichst reliabel sein sollte (Stier, 1999). Beobachterfehler können durch verschiedene Ursachen entstehen: mangelnde Leistungsbereitschaft / Leistungsfähigkeit des Beobachters, unpräzise definierte Kategorien, Erwartungshaltungen, Halo-Effekt, Interaktionen zwischen dem Beobachter und dem Geschehen oder Erinnerungsfehler (Huber, 1995). Boice (1983) führt die wichtigsten Massnahmen zur Vorbeugung von Beobachterfehlern auf:

- Seriöse Vorbereitung und Zeitplanung, d.h. theoriegeleitetes Beobachtungsschema erstellen, mit einer sinnvollen Segmentierung;
- Wenige Kategorien im Beobachtungsschema, die präzise, aber nicht zu eng formuliert sind;
- Unauffälliges Vorgehen bei der Beobachtung (solange ethisch vertretbar) und geeignete Objekte beobachten;
- Möglichst wenig Interaktion zwischen Beobachter und Beobachteten sowie richtige Auswahl und Training der Beobachter.

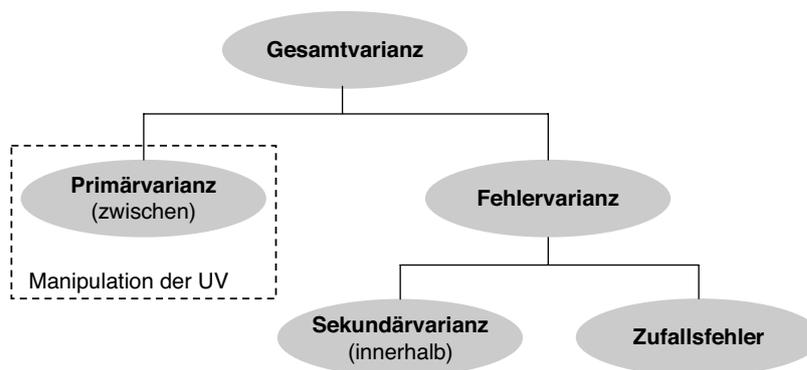
Jede Kategorie sollte sprachliche Einheiten unter dem Aspekt der semantischen Ähnlichkeit (Bedeutung) zusammenfassen. Daher sind die Kategorien zu operationalisieren, z.B. durch Vorgabe typischer Beispiele oder einer Stichwortliste (Stier, 1999). Beim Aufbau eines Beobachtungsschemas mit Kategorien gelten für eine Beobachtung dieselben Hinweise bezüglich Operationalisierung und Skalierung wie bei einem Fragebogendesign (Greve und Wentura, 1997). Kategorien können auch als grobes Raster vorliegen. Kromrey (2006) und Schnell et al. (1999) klassifizieren Beobachtungen in vier Grundtypen: (1) offene / verdeckte Beobachtung (wissen Personen, ob sie beobachtet werden?), (2) teilnehmende / nicht teilnehmende Beobachtung (nimmt der Beobachter am Geschehen aktiv teil?), (3) natürliche / künstliche Situation (Beobachtung im Feld oder im Labor?) und (4) systematische / unsystematische Beobachtung (liegt ein Kategoriensystem vor oder nicht?).

A.2.2 Experiment

Theorien zu beweisen, ist nicht das Ziel der Forschung (Sarris, 1999). Theorien können durch empirische Befunde lediglich gestützt werden: Es ist nicht auszuschliessen, dass eine bestätigte Theorie durch eine überlegenere zu ersetzen ist (Nagel, 1991). Ein solches Verständnis wissenschaftlicher Forschung kennzeichnet auch die Experimentalforschung (Hicks, 1993). Sie orientiert sich dabei an den Grundgedanken des kritischen Rationalismus (vgl. Kap. 1.5.1).

In einem Experiment ist die Primärvarianz, ausgelöst durch die Manipulation, von Interesse. Die Primärvarianz ist die Datenfluktuation, die allein auf die Manipulation der unabhängigen Variablen zurückzuführen ist. Um die Primärvarianz möglichst genau erfassen zu können, muss die Datenvarianz *zwischen* den beiden experimentellen Bedingungen von der Varianz *innerhalb* dieser Bedingung (Fehlervarianz) unterschieden werden (Bortz, 2006). Bei der inferenzstatistischen Überprüfung eines experimentellen Datensatzes werden diese beiden Varianzen zueinander in Beziehung gesetzt (vgl. Abb. A-1). Je höher die Primärvarianz gegenüber der Fehlervarianz, desto leichter lässt sich ein experimenteller Effekt nachweisen (Sarris und Reiss, 2005).

Abbildung A-1: Primär- und Fehlervarianz als Varianzquellen



Quelle: Eigene Darstellung in Anlehnung an Sarris und Reiss, 2005

Die Hauptergebnisse einer experimentellen Untersuchung werden als Durchschnittswerte (\bar{x}) ermittelt, da im Rahmen experimenteller Untersuchung von Verhaltensreaktionen naturgemäss Datenvarianz entsteht (Kirk, 1995; Hicks, 1993).

A.2.3 Nebenhypothesen

In Tabelle A-1 sind alle in dieser Arbeit experimentell getesteten Nebenhypothesen und die involvierten Variablen aufgeführt.

Tabelle A-1: Nebenhypothesen und involvierte Variablen

Nebenhypothesen: Interaktion (i) und Demografie (d)		Variablen
h_{3,i1}	Der Ticketpreis hat einen moderierenden Effekt auf die Kundensteuerung.	UV3 → UV1
h_{3,i2}	Der Ticketpreis hat einen moderierenden Effekt auf die Kommunikation.	UV3 → UV2
h_{4,i1}	Das ex ante Image hat einen moderierenden Effekt auf die Kundensteuerung.	UV4 → UV1
h_{4,i2}	Das ex ante Image hat einen moderierenden Effekt auf die Kommunikation.	UV4 → UV2
h_{d1a}	Je tiefer die Klassenzugehörigkeit (Produkt), umso tiefer ist die Akzeptanz des Kundensteuerungskonzepts.	-
h_{d1b}	Je tiefer die Klassenzugehörigkeit (Produkt), umso höher ist die Wahrscheinlichkeit für negative Verhaltensabsichten.	-
h_{d2a}	Je grösser die Flugerfahrung, umso grösser ist die Akzeptanz des Kundensteuerungskonzepts.	-
h_{d2b}	Je grösser die Flugerfahrung, umso tiefer ist die Wahrscheinlichkeit für negative Verhaltensabsichten.	-

Quelle: Eigene Darstellung

A.3 Grundzüge der statistischen Datenanalyse

In diesem Kapitel werden die theoretischen, statistischen Grundlagen zur Prüfung von Hypothesen und Wirkungsmodellen vorgestellt. Es werden nur die statistischen Instrumente erläutert, die zur Auswertung der vorliegenden Experimente eingesetzt werden. Für die Grundlagen zu den Wahrscheinlichkeitsverteilungen, insbesondere die Normal- und die *t*-Verteilung sowie Grundprinzipien zur Hypothesenformulierung wird auf einschlägige Literatur hingewiesen (z.B. Bortz, 2006; Kromrey, 2006; Hirsig, 2001). Nachdem in Kapitel A.3.1 Grundlagen zu den „Korrelationen“ erläutert werden, wird auf entscheidungsstatistische- (Kap. A.3.2) sowie varianzanalytische Verfahren (Kap. A.3.3) eingegangen. Da entscheidungsstatistische Verfahren über die Stichprobengrösse manipulierbar sind, werden in Kapitel A.3.4 die Berechnung der Effektgrösse und Richtwerte aus der Sozialforschung diskutiert.

A.3.1 Korrelationen

Die im Experiment untersuchten Merkmalsdimensionen stehen gemäss Wirkungsmodell untereinander in Beziehung. Im Modell haben unabhängige Variablen (UVn) eine gerichtete Beziehung auf die abhängigen Variablen (AVn). Korrelationen zeigen, inwieweit die Ausprägungsgrade von zwei Merkmalsdimensionen zusammenhängen. Eine geringe Ausprägung in der einen Merkmalsdimension führt zu einer geringen (oder höheren) Ausprägung in der anderen Merkmalsdimension (Cramer und Kamps, 2007).

Korrelation nach Pearson

Die „Enge“ eines stochastischen Zusammenhangs bei intervallskalierten Variablen wird mit den Masszahlen *Kovarianz* und mit der *Korrelation* beschrieben. Im Gegensatz zur Kovarianz weist das Mass der Korrelation einen wesentlichen Vorteil auf: Die Korrelation ist nicht abhängig von den verwendeten Massstäben (Skalierung der Variablen) und von den Streuungen in den Merkmalsverteilungen (Hirsig, 2001). Da in den Sozialwissenschaften die einer Skalierung zugrunde liegenden Massstäbe praktisch nie vorgegeben sind, wird die Korrelation zur Beschreibung eines stochastischen Zusammenhanges vorgezogen (Kirk, 1995). Sie ist von den Massstäben der verwendeten Skalen und Streuungen der Merkmalsausprägungen unabhängig. Der Korrelationskoeffizient errechnet sich, indem der Kovarianzwert in einen einheitlichen Wertebereich transformiert wird (Cramer und Kamps, 2007; Schira, 2003):

$$r = \frac{\text{cov}(x,y)}{s_x \cdot s_y}$$

$$(\text{wobei } \text{cov}(x,y) = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n})$$

Die Kovarianz $\text{cov}(x,y)$ der beiden Merkmalsausprägungen wird durch das Produkt der Standardabweichungen (s) der beiden Merkmalsverteilungen s_x und s_y dividiert. Damit wird sie in den Wertebereich zwischen -1,0 und +1,0 transformiert. Dem Vorzeichen kann entnommen werden, ob x und y mehrheitlich gleichsinnig oder mehrheitlich gegensinnig korreliert sind (Bortz, 2006). Wenn $r = 0$, dann sind x und y nicht korreliert. Die Berechnung des Korrelationskoeffizienten ist nur zulässig, wenn eine lineare Beziehung angenommen wird (Bamberg et al., 2007). Der so definierte Korrelationskoeffizient wird als *Pearson-Bravais-Korrelation* bezeichnet (Cramer und Kamps, 2007).

Punkt-biseriale Korrelation zwischen UV und AV

Die in der vorliegenden Untersuchung verwendeten unabhängigen Variablen haben zwei Ausprägungen (hoch, tief). Es handelt sich um ein dichotomes Merkmal. Da die (gemessene) abhängige Variable jeweils intervallskaliert ist, spricht man für den allfälligen stochastischen Zusammenhang zwischen UV und AV von *Punkt-biserialer Korrelation* (Hirsig, 2001). Die Berechnung der Punkt-biserialen Korrelation erfolgt in SPSS durch eine Produkt-Moment-Korrelation (als Definitionsgleichung, vgl. Bortz, 2006), wobei das dichotome Merkmal mit „1“ und „2“ codiert wird.

Multiple Regressionsanalysen

Die in dieser Untersuchung verwendete *multiple Korrelation* beschreibt die „Enge“ des stochastischen Zusammenhangs zwischen den prognostizierten Ausprägungen der AV und den tatsächlich beobachteten Ausprägungsgraden (Hirsig, 2002). Als *multiple Regression* wird eine algebraische Gleichung verstanden, die eine Vorhersage des Ausprägungsgrades der AV aus der UV ermöglicht (Bortz, 2006). Multiple Regressionsanalysen werden in der vorliegenden Arbeit mit dem Statistikprogramm SPSS durchgeführt.

Grösse des Korrelationskoeffizienten

Damit eine Korrelation signifikant ist, ist neben der Grösse des Unterschiedes auch die Grösse der Stichprobe entscheidend. Bei genügend grosser Stichprobe ist auch ein kleiner Unterschied in den Verteilungskennwerten signifikant (vgl. Schira, 2003). Je grösser der Umfang der Stichprobe, umso geringer muss eine Korrelation ausfallen, um signifikant zu sein. Die Mindestgrösse von r kann in der Tabelle für die Produkt-Moment-Korrelation bei ein- und zweiseitigen Hypothesen nachgesehen werden (Hirsig, 2001). In Tabelle A-2 wird dieser Zusammenhang verdeutlicht:

Tabelle A-2: Mindestgrössen von signifikanten Korrelationen

Freiheitsgrade	Wahrscheinlichkeit	
	5%	1%
<i>df</i>		
10	0.576	0.708
45	0.288	0.372
50	0.273	0.354
100	0.195	0.254
150	0.159	0.208
300	0.113	0.148

Quelle: Eigene Darstellung in Anlehnung an Hirsig, 2002

A.3.2 Entscheidungsstatistische Verfahren

Gerade im Bereich von Experimenten werden entscheidungsstatistische Fragestellungen relevant: sind die Unterschiede in den Stichprobenverteilungen zufällig entstanden oder durch Manipulation der unabhängigen Variablen? Die Hauptannahme bei der Formulierung der Nullhypothese ist, dass die beiden Stichproben je aus derselben Population ($H_0: \mu_1 = \mu_2$) stammen (Cramer und Kamps, 2007; Hirsig, 2001). Das logische Komplement zur Nullhypothese ist die Alternativhypothese (H_1). Sie ist im Gegensatz zur Nullhypothese statistisch nicht bearbeitbar und deckt einen Bereich ab (z.B. $\mu_1 \neq \mu_2$). Basierend auf Wahrscheinlichkeiten wird die Entscheidung getroffen, ob eine Nullhypothese falsifiziert oder nicht falsifiziert werden soll. Hierzu können verschieden strenge Massstäbe angelegt werden. Im Arbeitsbereich der Sozialwissenschaften haben sich die Irrtumswahrscheinlichkeiten (Signifikanzniveaus) von 5%, 1% und 0.1% etabliert (Hamlin, 2005; Hicks, 1993).

Im Experiment werden verschiedene Stichproben bei Probanden gezogen. Jede Stichprobe unterscheidet sich theoretisch durch die vorgenommene Manipulation der unabhängigen Variable. Basierend auf der Manipulation und unter der Annahme, dass beide Stichproben aus ein und derselben Population gezogen werden, wird bei der Hypothesenprüfung im Experiment auf Unterschiede bei den Mittelwerten der gemessenen Ausprägungen (AV) geachtet (Bamberg et al., 2007; Bortz, 2006; Diekmann, 1998).

Vergleich zweier Stichprobenmittelwerte aus unabhängiger Stichprobe

Statistische Verfahren zur Prüfung von Unterschiedshypothesen sind vom Skalenniveau abhängig (Cramer und Kamps, 2007; Hussy und Jain, 2002). Im Falle des Experiments basiert die Unterschiedshypothese auf dem Skalenniveau der gemessenen abhängigen Variablen. In der vorliegenden Untersuchung sind die Variablen intervallskaliert und die Stichproben unabhängig (keine Messwiederholung). Daher werden die Stichprobenmittelwerte mit dem t -Test für unabhängige Stichproben überprüft (Duller, 2007; Sarris und Reiss, 2005; Hirsig, 2001). Geprüft wird die Differenz der Stichprobenmittelwerte $(\bar{x}_1 - \bar{x}_2)$. Das arithmetische Mittel wird gewählt, weil (1) eine grosse Zuverlässigkeit verlangt wird, da \bar{x} , von einer Stichprobe zur anderen üblicherweise wenig schwankt, (2) andere Kennwerte wie bspw. die Varianz berechnet werden sollen und (3) weil angenommen wird, dass die Verteilung symmetrisch ist (Sarris und Reiss, 2005).

Um die Prüfverteilung bestimmen zu können, müssen die Populationsvarianzen geschätzt werden. Diese Schätzung hängt davon ab, ob die Varianzen der Population homogen oder heterogen sind (Hicks, 1993; Hirsig, 2001). Ob homogene oder heterogene Stichprobenvarianzen vorliegen, wird in SPSS anhand des *Levene-Tests der Varianzgleichheit* ermittelt (Duller, 2007). In SPSS wird bei $p \geq .05$ unter „*Equal variances assumed*“ von homogenen Varianzen ausgegangen (zufälliger Varianzunterschied), bei $p < .05$ von heterogenen Varianzen (Bamberg et al., 2007; Duller, 2007). Je nachdem, ob Varianzgleichheit vorliegt oder nicht („*Equal variances not assumed*“), muss für den t -Test der Mittelwertgleichheit mit einem anderen Schätzwert weitergefahren werden (Hirsig, 2001). Für homogene Varianzen gilt folgende Schätzung der Standardabweichung der Verteilung der Mittelwertsdifferenzen zweier Stichproben (Kirk, 1995):

$$\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = \hat{\sigma} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Für heterogene Varianzen gilt folgende Schätzung:

$$\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = \hat{\sigma} \cdot \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$$

Da die Prüfgrössen t -verteilt sind, stellt sich die Frage nach den Freiheitsgraden df , welche die Form der t -Verteilung festlegen. Dabei ist der Freiheitsgrad für homogene Varianzen definiert als $df = (n_1 + n_2 - 2)$ und für heterogene Varianzen als $df = (n_1 + n_2 - 2) / 2$ (Kirk, 1995; Hicks, 1993).

Für Mittelwertvergleiche *mehrerer Stichproben* wird in verschiedenen statistischen Untersuchungen ein t -Test (Zweistichprobenfall) mehrfach angewandt (Hirsig, 2002). Dies kann jedoch zu Pseudoeergebnissen (Artefakte) führen (Sarris und Reiss, 2005). Für Mittelwertvergleiche mehrerer Stichproben ist die in Kapitel A.3.3 beschriebene Varianzanalyse vorzuziehen (Sarris und Reiss, 2005).

Überprüfung der Signifikanz von Korrelationskoeffizienten

Auch Korrelationskoeffizienten können darauf geprüft werden, ob sie in einer Stichprobe zufällig entstanden sind (vgl. Kap. A.3.1). Bei der Überprüfung der Signifikanz von Korrelationskoeffizienten handelt es sich um ein univariates Prüfverfahren (Schira, 2003). Es wird geprüft, inwieweit ein in einer Stichprobe ermittelter Korrelationskoeffizient „zufällig“ entstanden sein kann. Die geprüfte Hypothese H_0 postuliert, dass die Korrelation zwischen den interessierenden Merkmalen in der Population gleich null ist (Hirsig, 2001; Kirk, 1995): $H_0: r = 0,0$. Für die Prüfgrösse bei $4 \leq n \leq 1'000$ wird die von R. A. Fisher 1918 entwickelte Formel verwendet (vgl. Hicks, 1993):

$$t = \frac{r}{\sqrt{1-r^2}} \cdot \sqrt{n-2}$$

(wobei r = beobachteter Korrelationskoeffizient)

In SPSS wird unter „Sig. (2-seitig)“ die zweiseitige Überschreitungswahrscheinlichkeit (Signifikanz) für den ermittelten Ausprägungsgrad des Korrelationskoeffizienten ausgegeben (Duller, 2007). Bei einem Korrelationskoeffizienten bspw. von $r = .67$ bedeutet „Signifikanz (2-seitig): .00“, dass mit einer Irrtumswahrscheinlichkeit von kleiner als 0,1% davon ausgegangen werden kann, dass die beiden untersuchten Merkmale mit $r = .67$ korrelieren und diese Korrelation nicht zufällig zustande gekommen ist.

Die Testung der Modellgüte nur aufgrund von Korrelationskoeffizienten ist allerdings fehlerhaft. Die Signifikanz von Korrelationen besagt nicht, dass ein zugrunde liegendes Modell den Daten genügt (Birnbaum, 1974). Gemäss Sarris (1999) könnte es sein, dass bei einer varianzanalytischen Modelltestung (vgl. Kap. A.3.3) statistisch signifikante Abweichungen zwischen den Modellwerten und den empirischen Daten nachgewiesen werden. Bei der Wahl einer ausschliesslich „korrelativen“ Modelltestung wird in der Regel ein Modell weit weniger widerlegt als im varianzanalytischen Fall. Sarris und Reiss (2005, S. 142) folgern: „Man sollte dieses inferenzstatistische Auswertungsverfahren [Verfahren der Varianzanalyse] generell als bestindizierte Technik zur Testung von zwei- oder mehrfaktoriellen Modellen verwenden.“

Teststärke

Im Zusammenhang mit Prüfverfahren ist die Teststärke (*Power*) ein wichtiges Mass. Sie ist definiert als $1-\beta$ (für β vgl. Kap. 6.2) und gibt die Wahrscheinlichkeit an, mit der eine korrekte Alternativhypothese H_1 angenommen wird (Sarris und Reiss, 2005; Hirsig, 2001; Kraemer und Thiemann, 1987). Es handelt sich also um die Wahrscheinlichkeit, einen tatsächlich vorhandenen Unterschied oder Zusammenhang zu entdecken. Bei einem nicht-signifikanten Ergebnis (Annahme der Nullhypothese) ist eine möglichst grosse Teststärke wichtig, weil nur damit die Richtigkeit der Nullhypothese bewiesen ist. Generell wird auch eine hohe Teststärke für signifikante Ergebnisse gefordert, wobei ein Kompromiss mit dem α -Fehler eingegangen werden muss. Als Faustregel wird dementsprechend eine Teststärke zwischen 0,7 und 0,9 verlangt (Kraemer und Thiemann, 1987).

Goodness of Fit Test

In der Regel wird bei prüfstatistischen Verfahren vorausgesetzt, dass das interessierende Merkmal in der Population wenigstens näherungsweise normalverteilt ist. Ist die Verteilung in der Population unbekannt, so kann die Verteilung in der Stichprobe mit einer Normalverteilung verglichen werden (Hirsig, 2001; Kirk, 1995). Weichen die Stichprobendaten nur zufällig von einer Normalverteilung ab, so kann davon ausgegangen werden, dass das Merkmal auch in der Population normalverteilt ist. Die Nullhypothese bei dieser Untersuchung lautet also: „Die empirische Verteilung der Stichprobe und die Normalverteilung unterscheiden sich nur zufällig.“

SPSS stellt für diese Untersuchung verschiedene Prüfverfahren zur Verfügung, darunter den *Kolmogorov-Smirnov Goodness of Fit-Test* (Duller, 2007; Johnson und Wichern, 2002). Zur Reduzierung des β -Fehler-Risikos (irrtümliche Annahme von H_0) wird dabei die Nullhypothese auch dann falsifiziert, wenn die Überschreitungswahrscheinlichkeit kleiner ist als 25% (Bamberg et al., 2007). Der *Kolmogorov-Smirnov-Test* ist jedoch von der Stichprobengröße abhängig. Bei grossen Stichproben ($n > 100$) werden bereits leichte Abweichungen von der Normalverteilung signifikant (Diehl und Arbinger, 2001). Alternativ können bei einer grossen Stichprobe Q-Q Plots (Quantil-Quantil Plots) gezeichnet werden. Diese zeigen die Abweichung der Normalverteilung grafisch auf. Das Q-Q Plot ist ein Streudiagramm, bei dem auf der x-Achse die beobachteten Quantile und auf der y-Achse die erwarteten Quantile abgetragen werden (Johnson und Wichern, 2002). Aufgrund der Robustheit der Tests in der vorliegenden Untersuchung (z.B. *F-Test* in ANOVA) genügt eine visuelle Kontrolle des Plots (Bortz, 2006; Roberts, 1999).

A.3.3 Varianzanalytische Verfahren in Wirkungsmodellen

Das in Kapitel A.3.2 besprochene entscheidungsstatistische Verfahren erlaubt den Nachweis der Unterschiedlichkeit von zwei Stichproben. Sollen aber die Auswirkungen mehrerer Stichproben-Beschreibungsmerkmale (Gruppen) simultan geprüft werden, ist dies nur mit einer Varianzanalyse möglich (Hirsig, 2002; Hussy und Jain, 2002). Sie ist eine Methode zum Vergleich von Stichprobenmittelwerten. Sarris (1999) und Birnbaum (1974) nennen die Varianzanalyse als die einschlägige Methode zur Modellprüfung bei Experimenten, vor allem bei zwei- oder mehrfaktoriellen Modellen. In SPSS heisst das Verfahren ANOVA (Analysis of Variance) und wird in der vorliegenden Untersuchung für den Manipulationscheck, die Interaktionseffekte sowie für die Prüfung der Hypothesen 8, 9 und 10 im Online-Experiment verwendet.

Univariate, mehrfaktorielle Varianzanalyse

In der vorliegenden Untersuchung werden die Auswirkungen mehrerer Faktoren (UV) auf eine abhängige Variable untersucht, ohne Messwiederholung. Es handelt sich um eine univariate, mehrfaktorielle Varianzanalyse. Die Analyse der Wechselwirkung zwischen den UV (UVA x UVB) ist zur Untersuchung des Interaktionseffekts geeignet (vgl. nächstes Kapitel). Um eine Varianzanalyse

ohne Messwiederholung durchzuführen, müssen zwei Voraussetzungen erfüllt sein (Bamberg et al., 2007; Hicks, 1993): (1) Die Ausprägungsgrade der AV in den Populationen (aus denen Stichproben stammen) müssen normalverteilt und intervallskaliert sein. Die Skalierung der UV spielt keine Rolle. (2) Die Varianz der Beobachtungen innerhalb der einzelnen Stichproben muss näherungsweise gleich sein. Das heisst, für alle Stichproben werden homogene Varianzen gefordert. In SPSS wird diese Voraussetzung mit dem Levene-Test überprüft (unter *Test der Homogenität der Varianzen*). Die Varianz der AV X in der Population wird in SPSS auf zwei unterschiedliche Arten geschätzt (Hussy und Jain, 2002; Hicks, 1993):

(1) Basierend auf der Varianz der Daten *innerhalb* der einzelnen Stichproben (Sekundärvarianz: systematischer Fehler). Die Werte der Varianz innerhalb der Stichproben werden in SPSS errechnet als *Summe-Abweichungsquadrat innerhalb der Gruppen* (SAQ_{in}). Sie ist in der SPSS ANOVA Tabelle bezeichnet als „Residuen“, unter der Spalte *Quadratsumme* (Duller, 2007).

(2) Die zweite Schätzung der Populationsvarianz stützt sich auf die Varianz *zwischen* den Stichprobenmittelwerten um den Gesamtmittelwert aller Daten (Primärvarianz, als Resultat der Manipulation). Die Werte der Varianz zwischen den Stichprobenmittelwerten wird in SPSS errechnet als *Summe-Abweichungsquadrat zwischen den Gruppen* (SAQ_{zw}). Sie ist in der SPSS ANOVA Tabelle bezeichnet als „Regression“, unter der Spalte *Quadratsumme* (Duller, 2007).

Ziel eines Experiments ist, die Primärvarianz zu maximieren und die Sekundärvarianz zu minimieren (Hussy und Jain, 2002). Eine Varianzanalyse basiert auf dem Vergleich dieser beiden Schätzwerte mit dem *F*-Test und folgender Prüfgrösse (Bortz, 2006):

$$F = \frac{SAQ_{zw}/df_{zw}}{SAQ_{in}/df_{in}}$$

(wobei bspw. SAQ_{zw}/df_{zw} = Schätzung der Populationsvarianz_{zw})

Die zentralen Kennwerte einer Varianzanalyse sind die *SAQ*-Werte (Summe der Abweichungsquadrate) und die ihnen zugeordneten Freiheitsgrade *df*. Führt eine Varianzanalyse zu einem signifikanten *F*-Wert ($F > 1,0$), so kann mit der zugehörigen Irrtumswahrscheinlichkeit *p* (5%, 1% oder 0,1%) geschlossen werden, dass die Ausprägungsgrade der UV mit den Ausprägungsgraden der AV X in einem überzufälligen Zusammenhang stehen (Schira, 2003). Bezüglich der Stichprobenmittelwerte bedeutet dies, dass mindestens einer der Vergleiche der Mittelwerte signifikant ist. Wird die Nullhypothese hingegen angenommen, dann steht die UV mit der AV in keinem statistisch nachweisbaren Zusammenhang (Hirsig, 2002; Hussy und Jain, 2002).

Neben der *Quadratsumme* (SAQ_{Gesamt}) sowie den Freiheitsgraden *df*, dem *F*-Wert und *Signifikanz* wird in der ANOVA auch das *Mittel der Quadrate* (Schätzwert der Populationsvarianz) aufgeführt (Duller, 2007; Hirsig, 2002). Die ANOVA in SPSS zeigt nicht auf, welche Gruppen sich unterscheiden, sondern nur, ob es Mittelwertsunterschiede zwischen den Gruppen gibt (Kirk, 1995).

Varianzanalyse zur Aufklärung von Interaktionseffekten

Zur Untersuchung der Wechselwirkung zwischen zwei UVn wird eine univariate, zweifaktorielle Varianzanalyse verwendet. Untersucht wird die Auswirkung von zwei UVn auf den Ausprägungsgrad der AV X (Bamberg et al., 2007; Hicks, 1993). Wird der Zusammenhang zwischen der Wechselwirkung von UV A und UV B und der AV X signifikant, so bedeutet dies, dass die Wirkung einer UV auf die AV jeweils vom Ausprägungsgrad der anderen UV abhängig ist (Hirsig, 2002; Hussy und Jain, 2002). Es besteht also eine Interaktion zwischen UV A und UV B bezüglich der AV X. Diese Interaktion kann auftreten, auch wenn die Haupteffekte (A, B) nicht signifikant sind.

Anhand von Interaktionsdiagrammen, welche die Mittelwerte von AV X pro UV A und UV B Ausprägung darstellen, wird die Wechselwirkung (Moderator-Effekt; AxB) ersichtlich (vgl. Sarris und Reiss, 2005; Hussy und Jain, 2002). Eine Nullinteraktion liegt vor, wenn die Auswirkungen einer UV auf allen Stufen der anderen UV gleichen. Die UV wirken somit unabhängig voneinander auf die AV. Die Linien im Interaktionsdiagramm verlaufen parallel (Hussy und Jain, 2002).

Bei einer Interaktion ist die Gesamtwirkung mehr als die Summe seiner Teile; die Einzelfaktoren wirken nicht-additiv zusammen (Sarris und Reiss, 2005). Es wird unterschieden zwischen ordinalen und disordinalen Interaktionen. Wenn eine Faktorstufe der UV der anderen unter allen möglichen Stufen des zweiten Faktors in der Ausprägung mindestens ebenbürtig oder überlegen ist, liegt eine ordinale Interaktion vor. Bei einer ordinalen Interaktion können die Haupteffekte sinnvoll aus der Varianzanalyse interpretiert werden (Hirsig, 2002; Hussy und Jain, 2002). Fallen die Unterschiede der Faktorstufen von B unter UV A_{tief} und unter UV A_{hoch} in unterschiedlicher Richtung aus, spricht man von disordinaler Interaktion (Sarris und Reiss, 2005). Bei einer disordinalen Interaktion können die Haupteffekte nicht mit der Varianzanalyse sinnvoll interpretiert werden.

A.3.4 Effektgrösse

Die Annahme oder Ablehnung einer Hypothese hängt vom Stichprobenmittelwert, den Populationsparametern (μ_0 und σ_0) sowie von der Stichprobengrösse n ab (Cramer und Kamps, 2007; Bortz, 2006). Mit grösser werdendem n wird auch der z -transformierte Stichprobenmittelwert ($z_{\bar{x}}$) grösser, womit die Übertretungswahrscheinlichkeit $p(z \geq z_{\bar{x}})$ abnimmt. Bei einer genügend grossen Stichprobe wird daher ein sehr kleiner Unterschied zwischen dem Stichprobenmittelwert \bar{x} und dem Populationsmittelwert μ_0 signifikant (Hirsig, 2001). Wenn eine kleine Stichprobengrösse gewählt wird, wird ein Gruppenvergleich selten signifikant, auch wenn in der Population ein grosser Unterschied vorhanden ist (Hussy und Jain, 2002).

Die Signifikanz als einziges Mass für die Beurteilung eines experimentellen Effekts ist daher nicht in allen Fällen geeignet (Jacobs, 2003). Hierzu sind Effektgrössenmasse geeignet, die im Anschluss an ein signifikantes Ereignis für die Daten der Untersuchung berechnet und für die Interpretation der Ergebnisse berücksichtigt werden (Diehl und Arbinger, 2001). Die Effektstärke normiert die Unterschiede zwischen den Versuchsgruppen auf die Streuung der Testwerte. Zur Beurteilung

des effektive nachgewiesenen Effektes (post-hoc) von zwei Stichprobenmittelwerten (unabhängige Stichproben), wird folgende Formel für das Differenzmass (d) verwendet (Jacobs, 2003):

$$d_{eff} = \frac{(\bar{x}_1 - \bar{x}_2)}{\hat{\sigma}}$$

(wobei für die Schätzung der Populationsvarianz ($\hat{\sigma}$) eine Schätzformel mit der Streuung der Stichproben verwendet wird, da die Streuung der Populationsvarianz in den meisten Fällen unbekannt ist: $s = \sqrt{\frac{s_1^2 + s_2^2}{2}}$)

Aus dieser Formel wird ersichtlich, dass die Effektgrösse mit grösseren Unterschieden in den experimentellen Bedingungen und kleinerer Streuung innerhalb der Bedingungen wächst. Die Teststärke und der β -Fehler hängen von der Effektgrösse ab, neben dem α -Fehler und dem Stichprobenumfang (Hussy und Jain, 2002). Fehlen die klassischen Kennwerte zur Berechnung der Effektstärke, können sie geschätzt werden (vgl. Westermann, 2000). Für den Fall, dass weder vorherige Untersuchungen noch inhaltliche Überlegungen die Grösse des aufzufindenden Effekts definieren, hat Cohen (1988) eine Klassifikation vorgeschlagen (vgl. Tab. A-3). Sie beruht auf Erfahrungswerten aus dem Bereich der Sozialwissenschaften (Bortz, 2006; Jacobs, 2003; Hirsig, 2001):

Tabelle A-3: Einteilung der Effektgrössen nach Cohen (1988)

Effektgrösse	d	r	r^2
klein	0.20	0.10	0.01
mittel	0.50	0.30	0.09
gross	0.80	0.50	0.25

Quelle: Cohen, 1988

Schwache Effekte sind zu erwarten, wenn ein neues Forschungsgebiet bearbeitet wird, über das wenig bekannt ist (Hussy und Jain, 2002). Die AV und die UV können in diesem Fall nicht optimal ausgewählt werden. Ebenfalls klein sind die Effektgrössen bei Untersuchungen, welche nicht im Labor stattfinden. *Mittlere Effekte* lassen sich gemäss Cohen (1988) bereits mit „blossem Auge“ erkennen. In den Sozialwissenschaften sind *starke Effekte* eher selten (Hussy und Jain, 2002). Ein aufzufindender Mindesteffekt sollte explizit festgelegt werden. Wird der Mindesteffekt nicht explizit festgelegt, so wird diese Festlegung implizit durch den Stichprobenumfang sowie die Festlegung des Signifikanzniveaus getroffen Cramer und Kamps (2007).

A.4 Fragebogen

A.4.1 Fragebogen E-Mail-Umfrage

Sehr geehrter Vielflieger

Ich freue mich, dass Sie an dieser E-Mail-Befragung teilnehmen. Das Ausfüllen des Fragebogens nimmt ca. 10 Minuten in Anspruch. Alle aus der Befragung gewonnenen Daten werden vertraulich behandelt und anonym ausgewertet.

Der Fragebogen betrifft „Bestrafungen von Flugreisenden wegen Nichteinhaltens vertraglicher Vereinbarungen zwischen Kunde und Airline“ (z.B. Übergepäckgebühr). Füllen Sie bitte Ihre Antworten direkt in dieses (Antwort-) E-Mail hinein. Bei Mehrfachauswahl ein [x] in das entsprechende Feld einfügen. Wenn Sie eine Frage nicht beantworten möchten, bitte Feld leer lassen.

Vielen Dank für Ihre Unterstützung und freundliche Grüsse,
János Heé

1) Was verstehen Sie als Passagier unter einer „Bestrafung von Kunden“, wenn Sie eine Flugreise antreten? Bitte mit eigenen Worten beschreiben:

2) Haben Sie in den letzten drei Jahren eine „Bestrafung“ resp. eine „Strafgebühr aufgrund nicht eingehaltener Richtlinie“ erlebt? Bitte mit eigenen Worten beschreiben:

3) Was war oder wäre Ihre Reaktion bei einer Bestrafung? (Mehrfachantworten möglich)

- a) Überraschung []
- b) Ärger / Frust []
- c) Beschwerde einreichen []
- d) Freunden von diesem negativen Ereignis erzählen []
- e) Künftig keine Produkte dieser Firma kaufen []
- f) Mein Verhalten anpassen []

4) Falls Sie mit Wut oder Ärger auf eine Strafe reagiert haben, worauf gründete Ihre Reaktion? Bitte nur beantworten, wenn Sie sich bereits einmal bestraft fühlten (Mehrfachantworten möglich).

- a) Höhe der Strafe / Gebühr war unverhältnismässig im Vergleich zum Ticketpreis []
- b) Ungleichbehandlung (andere Kunden wurden nicht bestraft) []
- c) Als guter Kunde sollte ich keine Strafe / Gebühr bezahlen müssen []
- d) Mangelhafte Kommunikation der Bedingungen (z.B. unverständliche Sprache) []

5) Wie fair empfinden Sie folgende Situationen bei einer Flugreise auf einer Skala von 1 (fair) bis 5 (unfair)? Bitte Zahl (1, 2, 3, 4, 5) neben Antwort hinschreiben (3 = indifferent).

- a) Kein Zugang zur Lounge möglich (kann für CHF 40 gekauft werden) []
- b) Keine Umbuchung möglich, aufgrund der Ticket-Tarifbedingungen []
- c) Nur ein Handgepäck erlaubt []
- d) Gebühr für Ticket-Neuerstellung, weil Ticket verloren []
- e) Kein Upgrade mit Meilen in die nächst höhere Klasse möglich []
- f) Meilen werden nicht angerechnet, da es sich um ein Spezialticket handelt []
- g) Stornierungs- / Umbuchungsgebühren wegen Ereignissen, die kontrollierbar waren (z.B. zu spät erschienen) []
- h) Stornierungs- / Umbuchungsgebühren wegen Ereignissen, die nicht kontrollierbar waren (z.B. Krankheit, Unfall) []
- i) Gebühr für Übergepäck []

A.4.2 Fragebogen Experiment I

Der folgende Fragebogen ist eine von insgesamt sechzehn verschiedenen Versionen (vgl. Kap. 5.1.3). Die Fragebogen unterscheiden sich bezüglich der Manipulation mindestens einer Faktorstufe. Der Fragebogen wurde ebenfalls in einer englischen Version erstellt (vgl. A.4.3).



Universität St. Gallen: Passagierbefragung zum Kundenverhalten

Sehr geehrter Passagier

Wir freuen uns, dass Sie an dieser Befragung teilnehmen. Sie wird im Rahmen einer Doktorarbeit an der Universität St. Gallen bei Prof. Thomas Bieger durchgeführt. Das Ausfüllen des Fragebogens nimmt ca. 10-15 Minuten in Anspruch.

Alle aus der Befragung gewonnenen Daten werden vertraulich behandelt und anonym ausgewertet. Für Auskünfte und Unterstützung stehen wir Ihnen gerne zur Verfügung.

Als Dankeschön für Ihre Teilnahme verlosen wir unter allen vollständig ausgefüllten Fragebögen:

- **1 Flug für zwei Personen ab Zürich, frei aus 40 Destinationen in Europa wählbar**
- **3 Gutscheine für Amazon im Wert von je CHF 100.-**

Falls Sie an der Verlosung teilnehmen möchten, geben Sie bitte am Ende der Studie Ihre E-Mail-Adresse an.

Vielen Dank für Ihre Unterstützung. Sie helfen damit, Einsichten in das Kundenverhalten bei Flugreisen zu gewinnen.

Mit freundlichen Grüßen

János Heé

Einführung**Was ist Ihr Geschlecht?** männlich weiblich**In welchem Jahr sind Sie geboren?**

<19 _____>

Wo leben Sie zur Zeit?

Land: _____

Was ist Ihre höchste abgeschlossene Ausbildung?

- Schulobligatorium
- Lehre / Berufsschule
- Mittelschule, Gymnasium / Handelsschule
- Meisterdiplom / Fachausweis
- Höhere Fachschule
- Fachhochschule / Universität
- andere

Nachfolgend wird Ihnen eine Situation beschrieben, die Sie so während der Online-Buchung und der Reise mit einer Fluggesellschaft erleben könnten. Wir bitten Sie, sich in die Situation hineinzuversetzen. Beziehen Sie Ihre Antworten ausschliesslich auf diese Situation und die geschilderten Eigenschaften. Wir bitten Sie, jeweils die Antwortmöglichkeit anzukreuzen, die Ihrem Gefühl am besten entspricht.

Bitte beantworten Sie alle Fragen sorgfältig und vollständig. Es gibt keine „richtigen“ oder „falschen“ Antworten, sondern nur Ihre persönliche Einschätzung der jeweiligen Situation.

Stellen Sie sich bitte folgende Ausgangssituation vor:

Sie planen eine Flugreise, um drei Wochen Ferien an Ihrer Traumdestination zu verbringen. Dieser Interkontinentalflug wird zehn Stunden dauern. Sie werden viel Gepäck dabei haben.

Da mehrere Airlines die Strecke bedienen, wählen Sie die Fluggesellschaft Y-Air. Sie wissen, dass bei Y-Air der Service an Bord und am Boden sehr gut ist. Die Flugzeugflotte entspricht dem neuesten Stand der Technik. Bei Kundenumfragen schneidet Y-Air regelmässig überdurchschnittlich ab. Ihr neues Meilenguthaben wird nach diesem Flug genügend hoch sein, um einen ähnlich weiten Flug bei der Y-Air einzulösen.

Sie werden die Reise in 3 Monaten antreten. Ihren Flug buchen Sie über die Internetseite der Y-Air. Zuerst tippen Sie allgemeine Daten wie Startort und Destination sowie das gewünschte Datum ein.

Sie erhalten einen Langstreckenflug in der Economy Class für **CHF 3'000.-** inkl. Taxen / Gebühren.

Bevor Sie die Buchung über das Internet abschliessen, klicken Sie auf einen gut sichtbaren Hinweis. Sie lesen die dort deutlich beschriebenen Tarifoptionen und Bedingungen Ihres Tickets.

Am Tag der Abreise finden Sie am Check-In folgende Situationen a) bis d) vor:

a) Der Check-In-Mitarbeiter macht Sie darauf aufmerksam, dass Sie aufgrund von Tarifbedingungen 5 Kilo Übergepäck haben. Sie müssen CHF 150.- im separaten Y-Air Büro bezahlen (nach 20 Kilo Freigeepäck kostet jedes zusätzliche Kilo CHF 30.-).

Bitte kreuzen Sie pro Frage jeweils eine Wahrscheinlichkeit auf der Skala von 1-6 an.

Bezahlen Sie die Gebühr?	1 2 3 4 5 6 Nein, sicher nicht <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> Ja, mit Sicherheit [] Ich möchte nicht antworten
Wenn Nein, warum nicht?	[] zu teuer [] kein Interesse (lasse Teil des Gepäcks da) [] andere _____

Wie reagieren Sie auf den Sachverhalt?

Ich werde mich bei der Y-Air wegen des Ereignisses beschweren.	1 2 3 4 5 6 Nein, sicher nicht <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> Ja, mit Sicherheit [] Ich möchte nicht antworten
Ich erzähle meinen Freunden von diesem Ereignis.	1 2 3 4 5 6 Nein, sicher nicht <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> Ja, mit Sicherheit [] Ich möchte nicht antworten
Meine Lust, wieder mit der Y-Air zu reisen ist nach diesem Ereignis...	1 2 3 4 5 6 gesunken <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> gestiegen [] Ich möchte nicht antworten

Wie fair empfinden Sie den Sachverhalt?

<input type="radio"/> sehr unfair <input type="radio"/> unfair <input type="radio"/> eher unfair <input type="radio"/> neutral <input type="radio"/> eher fair <input type="radio"/> fair <input type="radio"/> sehr fair

Was wäre Ihre emotionale Reaktion?

<input type="radio"/> wütend <input type="radio"/> unzufrieden <input type="radio"/> negativ überrascht <input type="radio"/> neutral <input type="radio"/> positiv überrascht <input type="radio"/> zufrieden <input type="radio"/> glücklich
--

Ich empfinde diesen Sachverhalt als...

... Zusatznutzen, den ich mir zusätzlich zu meiner Buchung kaufen kann (ich darf Übergepäck transportieren).	1 2 3 4 5 6 Nein, sicher nicht <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> Ja, mit Sicherheit [] Ich möchte nicht antworten
... negative Konsequenz, die mir die Airline auferlegt, um mein Verhalten als Passagier zu steuern (Richtlinie der Y-Air).	1 2 3 4 5 6 Nein, sicher nicht <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> Ja, mit Sicherheit [] Ich möchte nicht antworten

Für die unten aufgeführten Szenarien b, c und d wurden den Probanden dieselben Fragen wie auf der vorhergehenden Seite gestellt (bezüglich Meilen-Upgrade in Szenario d wurden einzelne Fragestellungen leicht angepasst).

b) Wegen der Tarifbedingungen haben Sie keinen Zugang zur Lounge (Verpflegung, Zeitungen und Sitzgelegenheit). Anrecht auf Boarding mit Priorität haben Sie auch nicht. Im separaten Y-Air Büro können Sie diese Optionen für CHF 40.- kaufen.

c) Für eine von Ihnen beantragte flexible Änderung des Rückreisedatums wird eine Gebühr von CHF 200.- verlangt (wegen der Tarifoptionen). Sie müssen für die Zahlung zum separaten Y-Air Büro gehen.

d) Mit der Hälfte aller von Ihnen bisher gesammelten Flugmeilen können Sie am Check-In einen Upgrade in die Business Class bezahlen. Trotz verfügbaren Sitzen kann die gewünschte Änderung aus organisatorischen Gründen bei Y-Air nicht getätigt werden.

Fragen zur Ausgangssituation (S. 3)

Bitte versetzen Sie sich zurück in das anfangs geschilderte Ausgangsszenario auf S. 3:

Sie empfinden das Image der Y-Air, basierend auf der Beschreibung der Ausgangssituation (S. 3) eher als:

mässig ------ gut

Die Y-Air unterscheidet zwischen günstigen und teuren Tickets für Langstreckenflüge. In der eingangs (S. 3) beschriebenen Buchung handelt es sich eher um ein:

günstiges Ticket --- teures Ticket

Auf die Tarifoptionen werden Sie bei der Online-Buchung der Y-Air (S. 3) folgendermassen aufmerksam gemacht:

unklar --- klar

auf unfaire Art --- auf faire Art

Abschliessende Fragen

Wohin fliegen Sie heute?

<input type="checkbox"/> Europa	<input type="checkbox"/> Nord-/Südamerika	<input type="checkbox"/> Naher Osten / Afrika	<input type="checkbox"/> Asien / Australien
---------------------------------	---	---	---

Haben Sie Ihren heutigen Flug online gebucht?

Ja Nein

Welche Fluggesellschaft benutzen Sie heute?

Airline: _____

In welcher Klasse fliegen Sie heute?

Economy Business First

Sind Sie Mitglied in einem Vielfliegerprogramm?

Nein
 Ja, mit Status:
 Einstiegsniveau (z.B. einfacher Teilnehmer)
 Vielflieger (z.B. Frequent Traveller)
 Gold Member (z.B. Senator)
 Höchstes Niveau (z.B. HON Circle Member)

Wie oft sind Sie im 2006 geschäftlich geflogen?

- Genau <____> mal.
- Ich weiss nicht mehr wie oft, so zwischen <____> und <____> mal.

In der Funktion als:
 Angestellter (auf Kosten der Firma reisend)
 Unabhängiger Geschäftsmann (auf eigene Kosten reisend)
 Konferenzbesucher
 Sonstiges: _____

Wie oft sind Sie im 2006 privat geflogen?

- Genau <____> mal.
- Ich weiss nicht mehr wie oft, so zwischen <____> und <____> mal.

Vorwiegend um:
 Familie und Freunde zu treffen
 Urlaub, Freizeit zu verbringen
 Sonstiges: _____

Welche Gedanken hatten Sie beim Ausfüllen dieses Fragebogens?

Wettbewerb

Falls Sie am Wettbewerb teilnehmen möchten, schreiben Sie bitte Ihre E-Mail-Adresse oder Telefonnummer auf:

Diese Angabe wird nur für die Gewinnbenachrichtigung benötigt. Sie wird selbstverständlich getrennt von den Umfragedaten erfasst und vertraulich behandelt.

Der Rechtsweg ist ausgeschlossen und eine Barauszahlung der Gewinne nicht möglich.

Vielen Dank für Ihre Teilnahme.

A.4.3 Fragebogen Experiment II

Der Online-Fragebogen des zweiten Experiments steht auf:

<http://www.yoshi-air.ch/fragebogen> zu Testzwecken zur Verfügung.

Legende zur Beobachtung (A.5.1): 8Q = Onur Air, AB = Air Berlin, BA = British Airways, JL = Japan Airlines, LH = Lufthansa, LX = Swiss, QR = Qatar Airways, TG = Thai Airways, WK = Edelweiss Air, e = Economy Class, b = Business Class, g = Geschäftsreisender, f = Ferienreisender, et = Ethnic Travel, p = Privatreisender (allgemein), m = Mittelstrecke, k = Kurzstrecke, l = Langstrecke;

Zusatzinformation: Anzahl beobachtete Zwischenfälle = 42 Gruppen (82 Personen), Anzahl beobachtete Check-In (Gruppen) total = 478

A.5.2 Inhaltsanalyse des Online-Buchungsprozesses von Airlines

	GM	Strecke ¹	Ticketpreis (Economy ²)	Komm. Bedingungen ³	Preis der Strafe ⁴	Steuerung Anreize ⁵	Preis der Option ⁶
Air Asia	LCC	Manila - Kuala Lumpur	1: CHF 160 -	klar	CHF 8 - 25	Nein	VIP Boarding: CHF 7
Air France	NC	Paris - Bangkok	1: CHF 1'577 2: CHF 6'760	unklar	CHF 5 - 50	Nein	-
American Airlines	NC	New York - Zürich	1: CHF 1'437 2: CHF 3'215	unklar	CHF 55 (9 kg flat fee)	Nein	-
British Airways	NC	London - Bangkok	1: CHF 1'976 2: CHF 3'047	unklar	CHF 55 (9 kg flat fee)	Nein	-
Delta	NC	Atlanta - Frankfurt	1: CHF 1'576 2: CHF 3'600	klar	CHF 55 bis 9 kg 110.- bis 22 kg	teilweise	Lounge: CHF 28
Edelweiss Air	C	Zürich - Puerto Plata	1: CHF 1'657 (25.4 - 2.5.08)	unklar	CHF 10 - 15	Nein	-
LTU / Air Berlin	C; LCC	Düsseldorf - Cancun	1: CHF 1'780 2: CHF 2'630	unklar	CHF 16 - 33	Nein	-
Regional Express	R	Sidney - Lismore	1: CHF 183 2: CHF 606	unklar	CHF 3	Nein	-
Ryanair	LCC	London - Marrakesch	1: CHF 141 -	klar	CHF 13	teilweise	VIP Boarding: CHF 9
Swiss	N	Zürich - Bangkok	1: CHF 1'703 2: CHF 2'511	klar	CHF 10 - 70	teilweise	Umbuchen: CHF 200
Thomson-fly	C	London - Cancun	1: CHF 1'330 -	klar	Übergepäck als Anreiz	Ja	Übergepäck: CHF 110 (5 kg); Essen an Bord: CHF 22
US Airways Express	R	Phoenix - Boston	1: CHF 761 2: CHF 1'970	klar	55.- bis 9 kg 90.- bis 22 kg	Nein	-

Legende GM = Geschäftsmodell, C = Charter Airline, LCC = Low Cost Carrier, NC = Network Carrier, R = Regional Airline.

¹ = Wenn die Airline keine Langstreckenflüge anbietet, wird die längste bediente Strecke angezeigt. Als Abflugort wird, wenn möglich, der Heimatflughafen der Airline gewählt.

² = Economy Class Ticket: Hinflug 5. Juli, Rückflug 19. Juli 2008. Preis 1: ohne Flexibilität, Preis 2: mit Flexibilität. Es werden immer die günstigsten Tarife gewählt (ohne Flughafentaxe). Währungsumrechnung am Untersuchungstag mit <http://finance.yahoo.com/currency>.

³ = Kommunikation (Komm.) unklar, wenn mehr als 3 Klicks nötig sind od. unstrukturierte Darstellung der Bedingungen.

⁴ = Bestrafung: Übergepäck-Gebühr, Angaben pro Kilo Übergepäck (Anzahl Gepäckstücke nicht berücksichtigt).

⁵ = Anreiz-Optionen während des Buchungsprozesses: ja, nein, teilweise (wenn nur eine Option angeboten).

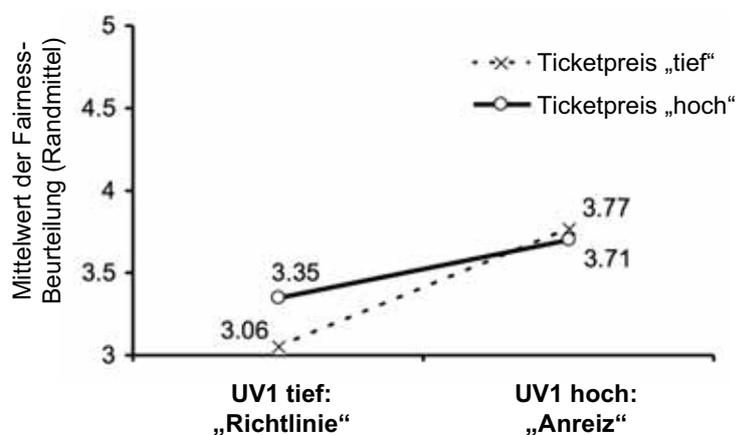
⁶ = Anreiz-Option mit Preis.

A.5.3 Forschungsergebnisse Experiment I

Interaktion zwischen Ticketpreis und Kundensteuerung ($h3_{i1}$)

Der nicht-signifikante p -Wert in Szenario A liegt bei .25 ($F = 1.34$, $df = 326$, $n_{tief} = 151$, $n_{hoch} = 178$). Trotz dieser Nicht-Signifikanz zeigt die grafische Analyse (vgl. Abb. A-2) eine Kreuzung der Geraden (disordinale Interaktion). Die Kreuzung der Geraden steht jedoch in einem überzufälligen Zusammenhang und bestätigt die Interaktion nicht. In den Szenarien B und C wird die Nullhypothese nicht falsifiziert (p (B) = .33; p (C) = .61). Szenario D hingegen weist eine signifikante Interaktion zwischen „Ticketpreis“ und „Kundensteuerung“ auf (vgl. Kap. 6.4.2).

Abbildung A-2: Interaktion zwischen Ticketpreis und Kundensteuerung (Szenario A)



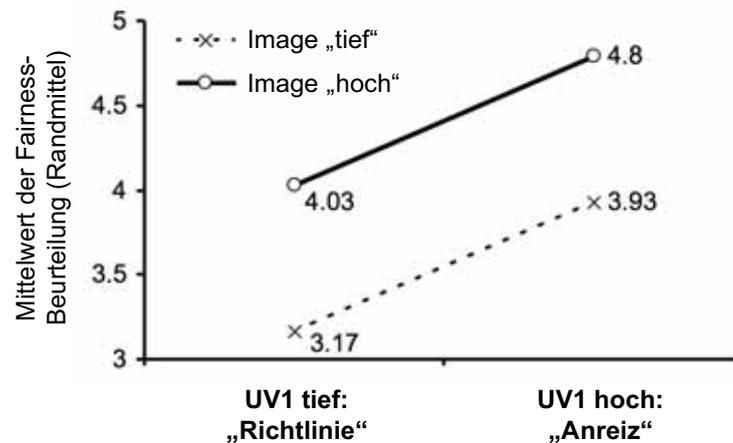
Quelle: Eigene Darstellung

Interaktion zwischen Ticketpreis und Kommunikation ($h3_{i2}$)

Für Szenario D wird auf dem Signifikanzniveau $p < 5\%$ eine disordinale Interaktion zwischen der UV3 (Ticketpreis) und der UV2 (Kommunikation) bestätigt (vgl. Kap. 6.4.2). Obwohl in Szenario A der p -Wert .23 ist, könnte aufgrund der grafischen Analyse eine disordinale Interaktion vermutet werden (vgl. Abb. A-2). Mit einer hohen Wahrscheinlichkeit von 23% ist dieser Zusammenhang jedoch zufällig entstanden, weshalb die Nullhypothese nicht falsifiziert wird (Szenario A: $F = 1.29$, $df = 327$, $n_{tief} = 149$, $n_{hoch} = 182$). In den Szenarien B und C besteht kein Interaktionseffekt, da die Fehlerwahrscheinlichkeit hoch ist (p (B) = .40, p (C) = .32).

Interaktion zwischen ex ante Image und Kundensteuerung ($h4_{i1}$)

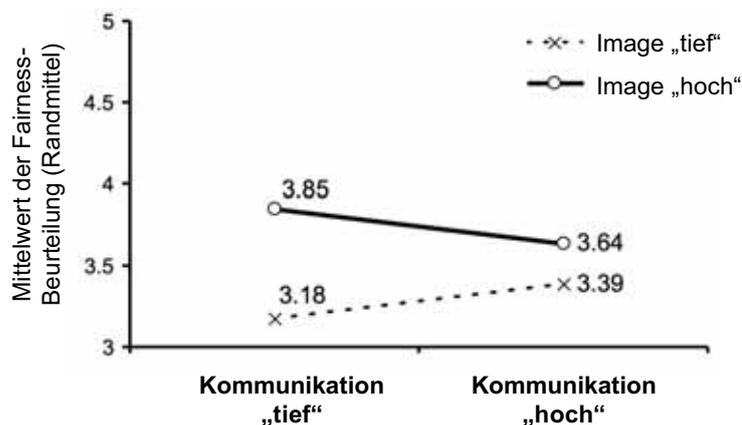
Kein Szenario bestätigt die Alternativhypothese für eine Interaktion zwischen *ex post Image* und „Kundensteuerung“ ($.97 > p_{A,B,C,D} > .15$, $df = 289$, $n_{tief} = 140$, $n_{hoch} = 152$). Die grafische Auswertung zeigt, dass eine Nullinteraktion vorliegt. Im Liniendiagramm in Abbildung A-3 verlaufen alle Linien parallel. Die beiden unabhängigen Variablen „Image“ und „Kundensteuerung“ wirken unabhängig voneinander auf die AV *Fairness-Beurteilung*.

Abbildung A-3: Interaktion zwischen Image und Kundensteuerung (Szenario B)

Quelle: Eigene Darstellung

Interaktion zwischen ex ante Image und Kommunikation (h4_{i2})

Mit dem tiefsten p -Wert von .19 in Szenario A wird die Nullhypothese in allen Szenarien nicht falsifiziert. Die grafische Analyse zeigt lediglich in Szenario A einen nicht-signifikanten, schwachen Moderatoreffekt (ordinale Interaktion, vgl. Abb. A-4). *Ex ante Image* ist somit keine Moderatorvariable bezüglich der UV „Kommunikation“.

Abbildung A-4: Interaktion zwischen Image und Kommunikation (Szenario A)

Quelle: Eigene Darstellung

Qualitative Auswertung der offenen Frage im Flughafen-Experiment

Die Probanden des Flughafen-Experiments konnten einzelne Szenarien kommentieren oder allgemein zum Fragebogen Gedanken äussern. Mit der qualitativen Auswertung (vgl. Tab. A-4) sind die Kommentare eingeteilt, je nach Szenario und manipulierter Variable „Kundensteuerung“ (d.h., Offline- oder Online-Gruppe).

Tabelle A-4: Qualitative Auswertung Flughafen-Umfrage

	Offline-Gruppe „Richtlinie mit Bestrafung“	Online-Gruppe „Anreiz mit Option“
Allgemein	<ul style="list-style-type: none"> – „Wurde zu schlecht und nicht transparent genug im Internet kommuniziert (versteckte Kosten).“ – „Mich ärgert, wenn es zu viele, teure Sonderoptionen gibt.“ – “The separate office would be an inconvenience whereby I may not pay the fee.” – “Read carefully these air-conditions and specific requests cost always a lot.” 	<ul style="list-style-type: none"> – „Bezahle es, falls Tarifbedingungen bei Buchung klar definiert waren.“ – „Mich ärgern die grossen Differenzen bei den Sitzplatzkategorien.“
Übergepäck	<ul style="list-style-type: none"> – „Ich habe nie mehr als 20kg.“ – „Sollte inbegriffen sein, bei diesem teuren Ticket.“ – „150 kg schwere Passagiere dürfen gleich viel mitnehmen wie ich mit meinen 53 kg. Unfair.“ – “Gives me an incentive to pack light.” 	<ul style="list-style-type: none"> – „Ich kann bei zuviel Übergepäck noch am Check-In zahlen.“ – „Man weiss, dass man nach 20kg bezahlen muss. Ich kenne jedoch die Kosten nicht.“ – “According to the new regulations limiting the hand luggage number they should allow such an excess in our luggage.”
Lounge-Zugang	<ul style="list-style-type: none"> – “I am a frequent flyer with access to alternative air lounges. This lounge-access is too expensive.” 	<ul style="list-style-type: none"> – „Lounge sollte allen zugänglich sein.“
Flexibles Ticket	<ul style="list-style-type: none"> – „Ferien sind Teil der Jahresplanung, Rückflug ist ein Fixdatum.“ 	<ul style="list-style-type: none"> – „Diese Option sollte gross und deutlich zu sehen sein.“ – „Bei der Buchung muss auch das kleinste Detail gelesen werden und bei zweideutigem Satzverständnis nachgefragt werden.“
Meilen-Upgrade	<ul style="list-style-type: none"> – „Lieber einen weiteren Flug.“ – „Da ich oft interkontinental fliege, wäre dieses Angebot positiv.“ 	

Quelle: Eigene Darstellung

A.5.4 Forschungsergebnisse Experiment II

Kundensteuerung und Fairness-Beurteilung (Replikation h1)

1. Formulierung Hypothese 1 (h1)

h1 modelltheoretisch formuliert: Wenn die Faktorstufe von UV1 (Kundensteuerung) tief (Richtlinie) ist, dann folgt eine tiefe Ausprägung der AV3, Fairness-Beurteilung (Beurteilung als unfair).

2. Stichprobengrösse

In beiden Szenarien B und D werden für h1 nur die Stichproben verwendet, bei denen der Manipulationscheck bezüglich Kundensteuerung funktioniert hat. In Szenario B sind dies 145 Probanden für die Faktorstufe „Richtlinie“ und 245 Probanden für die Faktorstufe „Anreiz“ (tot. in B: 390). In Szenario D können 295 Probanden der Stufe „Richtlinie“ zugeordnet werden und 275 der Stufe „Anreiz“ (tot. in D: 570). Wie aus Tabelle A-5 ersichtlich, wählen bei beiden Faktorstufen über 37% der Probanden beim Manipulationscheck „neutral“. Es gibt keine *Missings*.

Tabelle A-5: Manipulationscheck für Kundensteuerung (Online-Umfrage)

h1 (Online): Manipulations- check		M.c. funktioniert		M.c. nicht funktioniert		M.c. neutral		Total
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
B	Richtlinie	145		151		225		521
	Anreiz	245		96		218		559
	Total	390	36	247	23	443	41	1'080
D	Richtlinie	295		33		193		521
	Anreiz	275		66		218		559
	Total	570	53	99	9	411	38	1'080

Legende: M.c. = Manipulationscheck

Quelle: Eigene Berechnung mit SPSS

Die Auswertung der Flughafen-Umfrage zeigt, dass zwischen den Manipulationen „funktioniert“ und „nicht funktioniert“ ein signifikanter Unterschied besteht (vgl. Kap. 6.3.1). Mit dem *t*-Test wird daher bei h1 in der Online-Umfrage verglichen, ob für den Manipulationscheck UV1 „funktioniert“ ein Unterschied zu den „neutralen“ Fragebogen besteht. Die Nullhypothese lautet: „Zwischen den Gruppen Manipulationscheck UV1 „funktioniert“ und „neutral“ gibt es in Szenario B keine Unterschiede“. Die Nullhypothese wird auf dem $p < 0.1\%$ Signifikanzniveau abgelehnt ($t = -12.83$, $df = 378$, $n = 390$, $p = .00$). Für Szenario B dürfen nur die Fragebogen verwendet werden, die den Manipulationscheck bestanden haben. Es besteht ein Unterschied zwischen dem Mittelwert der Fairness-Beurteilung in Szenario B, je nachdem, ob Probanden die Manipulation verstehen oder „neutral“ wählen.

In Szenario D werden unter der Faktorstufe „Anreiz“ sowie „Richtlinie“ nur diejenigen Fragebogen verwendet, bei denen der Manipulationscheck funktioniert hat. Unter der Faktorstufe „Richtlinie“ bspw. ist $\bar{x}_{Mf} = 1.10$ ($s = 0.96$) und $\bar{x}_{Mn} = 2.58$ ($s = 1.16$). Die Mittelwertdifferenz der AV3 zwischen den Gruppen „Manipulationscheck funktioniert“ und „neutral“ ist nicht zufällig entstanden ($t = -14.69$, $df = 566$, $n = 570$, $p = .00$).

3. Hypothesenprüfung (Mittelwertsunterschiede)

Die Beurteilung der Fairness ändert sich in Szenario B vom Stichprobenmittelwert $\bar{x}_{KR} = 1.66$ (\bar{x}_{KR} (D) = 1.10) bei der Richtlinien- zu $\bar{x}_{KA} = 4.03$ (\bar{x}_{KA} (D) = 4.18) bei der Anreiz-Steuerung (vgl. Tab. A-6). Der *t*-Test für Mittelwertgleiche unabhängiger Stichproben ergibt, dass die Nullhypothese in beiden Szenarien falsifiziert wird. Die beobachteten Unterschiede pro Szenario B oder D können mit einer Irrtumswahrscheinlichkeit von $p < 0.1\%$ - nicht zufällig entstanden sein.

Tabelle A-6: Auswirkung der Kundensteuerung auf Fairness-Beurteilung (Online-Umfrage)

h1 (Online): UV1 → AV3		Kundensteuerung		t-Test (2-seitig)			Effekt
		Richtlinie	Anreiz	df	t	Sig.	d
B	n	145	245				
	Mittelwert	1.66	4.03	261	-18.17	.00	1.96
	Stand.abw.	1.29	1.13				
D	n	295	275				
	Mittelwert	1.10	4.18	535	-34.72	.00	2.93
	Stand.abw.	0.96	1.14				

Quelle: Eigene Berechnung mit SPSS und G*Power

4. Effektgrösse und Teststärke

Das von der Stichprobe unabhängige Mass der Effektgrösse zeigt für beide Szenarien einen ausgeprägten Effekt zwischen den Faktorstufen „Richtlinie“ und „Anreiz“. In Szenario B beträgt das Differenzmass $d = 1.96$ und in Szenario D beträgt der starke Effekt $d = 2.93$. Die Teststärke ($1 - \beta$) beträgt 1 in beiden Szenarien.

Kommunikation und Fairness-Beurteilung (Replikation h2)

1. Formulierung Hypothese 2 (h2)

h2 modelltheoretisch formuliert: Wenn die Faktorstufe von UV2 hoch ist (klare Kommunikation), dann folgt eine hohe Ausprägung der AV3 (Beurteilung als fair).

2. Stichprobengrösse

Unter beiden Faktorstufen Kommunikation „unklar“ sowie „klar“ wird mit dem t -Test für Szenario B bestätigt, dass zwischen den Gruppen Manipulationscheck „funktioniert“ und „neutral“ signifikante Unterschiede bestehen ($t = -2.72$, $df = 756$, $n = 758$, $p = .01$). In Szenario B ist die Manipulation der Kommunikation (UV2) bei 400 Probanden erfolgreich (Faktorstufe „unklar“ $n = 128$, „klar“ $n = 272$, Faktorstufe).

In Szenario D wird die Nullhypothese für beide Faktorstufen nicht falsifiziert: Zwischen den Probanden, welche den Manipulationscheck der Kommunikation verstanden haben, und den Probanden, welche „neutral“ wählten, besteht kein Unterschied ($t = -1.21$, $df = 756$, $n = 758$, $p = .23$). Die Mittelwerte bspw. unter der Faktorstufe Kommunikation „klar“ weisen in Szenario D mit $\bar{x}_{Mf} = 3.13$ ($s = 1.68$) und $\bar{x}_{Mn} = 2.90$ ($s = 1.69$) eine geringe Differenz auf. Für Szenario D resultieren somit für beide Faktorstufen 758 Probanden für h2 (Faktorstufe „unklar“ $n = 314$, „klar“ $n = 444$).

3. Hypothesenprüfung (Mittelwertsunterschiede)

Bei der Interpretation des Mittelwertsunterschiedes zwischen den Faktorstufen Kommunikation „unklar“ und „klar“ darf sowohl in Szenario B als auch in D der Zufall mit einer Irrtumswahrscheinlichkeit kleiner als 0.01% ausgeschlossen werden (vgl. Tab. A-7).

Tabelle A-7: Auswirkung der Kommunikation auf Fairness-Beurteilung (Online-Umfrage)

h2 (Online): UV2 → AV3		Kommunikation		t-Test (2-seitig)			Effekt
		unklar	klar	df	t	Sig.	d
B	<i>n</i>	128	272				
	Mittelwert	2.59	3.78	398	-7.96	.00	0.84
	Stand.abw.	1.48	1.35				
D	<i>n</i>	314	444				
	Mittelwert	2.10	3.04	734	-8.31	.00	0.60
	Stand.abw.	1.42	1.68				

Quelle: Eigene Berechnung mit SPSS und G*Power

4. Effektgrösse und Teststärke

Mit einem Differenzmass von $d = 0.84$ wird in Szenario B ein starker Effekt nachgewiesen. Der Effekt in Szenario D ist von mittlerer Ausprägung, mit einem Wert von $d = 0.60$: Der Mittelwertsunterschied zwischen Kommunikation „tief“ und „hoch“ ist mittelstark ausgeprägt. In beiden Szenarien ist die Teststärke mit einem Wert von $(1 - \beta) = .99$ hoch.

Akzeptanz und negative Verhaltensabsichten (Replikation h7)

1. Formulierung Hypothese 7 (h7)

Die inhaltlich-theoretische Formulierung von h7 in der Online-Untersuchung entspricht h7 der Flughafen-Untersuchung (vgl. Kap. 6.3.7).

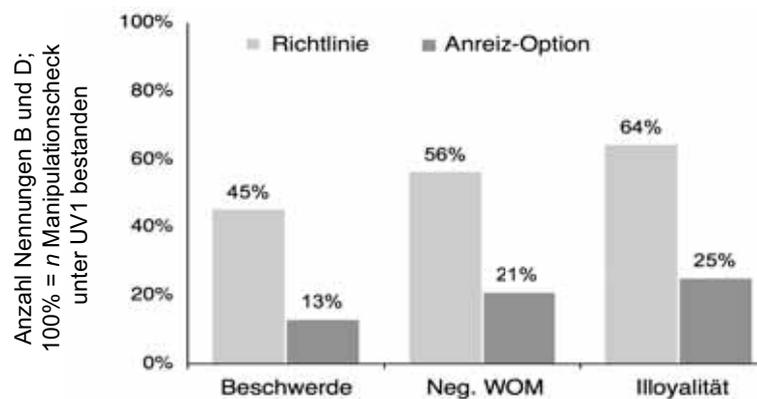
Die Skalierung von AV2 (neg. Verhaltensabsichten) im Online-Fragebogen lässt keine Mittelwertbildung zu, da mehrfache Nennungen in einer Nominalskala möglich sind: „Beschwerde“ (AV2_1), „neg. WOM“ (AV2_2), „Illoyalität“ (AV2_3) und „keine Antwort“. Daher kann h7, im Gegensatz zur Flughafen-Umfrage, nicht mit dem *t*-Test oder einer Varianzanalyse ausgewertet werden. Eine Kreuztabelle ermöglicht die Auswertung. Sie zeigt pro Kategorie der AV1-Intervallskala die Anzahl Nennungen unter AV 2_1-3 auf .

2. Stichprobengrösse

Die Voraussetzung, dass negative Verhaltensabsichten nur im Falle von Richtlinien entstehen (vgl. h7), wird im Flughafen-Experiment für die Bestimmung der Stichprobengrösse implizit angenommen. Zur Überprüfung dieser Annahme werden in der Online-Untersuchung die einzelnen Verhaltensabsichten unter der UV1 (Kundensteuerung) Faktorstufe „Richtlinie“ und „Anreiz“ gegenübergestellt.

In Szenario B äussern unter der Faktorstufe „Richtlinie“ deutlich mehr Probanden negative Verhaltensabsichten als unter der Faktorstufe Anreiz. Auch für Szenario D gilt, dass negative Verhaltensabsichten vor allem bei einer Richtlinie ausgelöst werden (vgl. Abb. A-5). Für die Prüfung von h7 werden daher nur Probanden der Faktorstufe „Richtlinie“ analysiert.

Abbildung A-5: Negative Verhaltensabsichten unter Richtlinie und Anreiz



Quelle: Eigene Darstellung

Der *t*-Test bestätigt für die Faktorstufe „Richtlinie“ ($n = 521$), dass zwischen den Gruppen Manipulationscheck „bestanden“ und „neutral“ signifikante Unterschiede im Antwortverhalten bestehen (vgl. h1). Für h7, Szenario B werden somit 145 Fragebogen untersucht, in Szenario D ist $n = 295$ (vgl. Tab. A-5).

3. Hypothesenprüfung Szenario B (Lounge-Zugang)

Von den Probanden, die den Manipulationscheck unter Kundensteuerung „Richtlinie“ bestehen, werden folgende negative Verhaltensabsichten geäußert (Mehrfachnennungen möglich, $n = 145$):

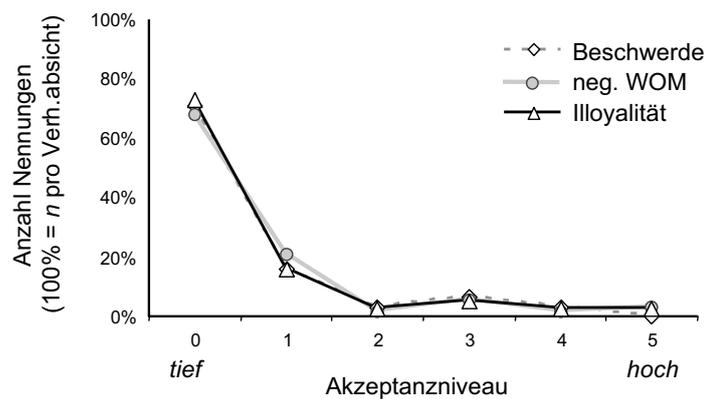
- 21 Prozent würden sich beschweren („werde Y-Air kontaktieren, um mich zu beschweren“);
- 49 Prozent „erzählen Freunden von diesem Vorfall“ und
- 58 Prozent verhalten sich illoyal („Meine Lust, wieder mit Y-Air zu reisen, ist nach diesem Ereignis gesunken“).
- In keine der oben genannten Kategorien fallen 17 Prozent der Probanden („Ich reagiere nicht wie hier beschrieben“).

Zur Überprüfung der Hypothese 7 wird in der Kreuztabelle die Summe der Anzahl Nennungen in den Kategorien „sicher nicht“ und „sehr tiefe Akzeptanz“ gebildet. Diese Summe wird prozentual mit der totalen Anzahl Nennungen pro negative Verhaltensabsicht (AV2_1-3) verglichen. Somit wird absolut und relativ untersucht, ob h7 zutrifft.

87 Prozent der 31 Probanden, die sich beschwerten, ordnen ihren Akzeptanzlevel den zwei tiefsten Akzeptanz-Kategorien zu. Von 71 Probanden, die angeben, dass sie den Freunden vom Vorfall erzählen werden, liegen 89 Prozent in den untersten beiden Akzeptanz-Kategorien. Bei der Illoyalität sind es sogar 88 Prozent, die auf den tiefsten Akzeptanzniveaus liegen (von $n = 84$).

Hypothese 7 wird für Szenario B bestätigt, da bei jeder negativen Verhaltensabsicht über 86% der Probanden in den zwei tiefsten Akzeptanzkategorien (von total sechs) liegen (vgl. Abb. A-6).

Abbildung A-6: Akzeptanz und negative Verhaltensabsichten (Szenario B)



Quelle: Eigene Darstellung

4. Hypothesenprüfung Szenario D (Meilen-Upgrade)

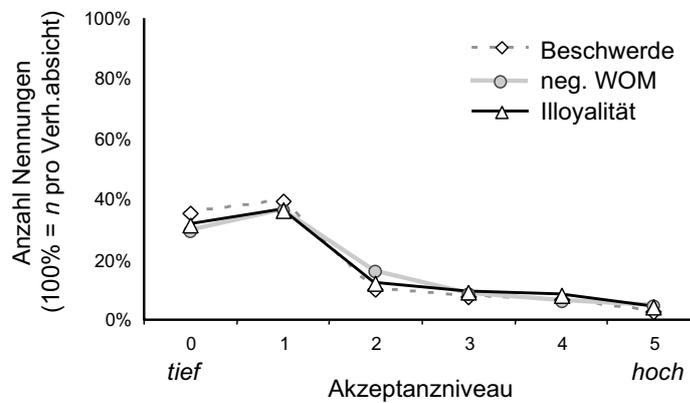
Von den 295 Probanden, die den Manipulationscheck bestanden haben, würden im Falle einer Bestrafung (Mehrfachnennungen möglich):

- 57 Prozent sich beschwerten;
- 60 Prozent ihren Freunden vom Vorfall erzählen und
- 66 Prozent sich illoyal verhalten.
- In keine der oben genannten Kategorien fallen 5 Prozent der Probanden („Ich reagiere nicht wie hier beschrieben“).

In Szenario D ist bei einer Richtlinie die Neigung zu negativen Verhaltensabsichten relativ höher als in Szenario B. Auch in Szenario D liegen die negativen Verhaltensabsichten bei den Probanden mit den tiefsten Akzeptanzniveaus, wie h7 vorhersagt. Von den 167 Probanden, die sich beschwerten würden, liegen 74 Prozent auf dem tiefsten und sehr tiefen Akzeptanzniveau, bei neg. WOM sind es 65 Prozent (von $n = 176$) und bei Illoyalität 67 Prozent (von $n = 196$). Nur ein Proband auf dem tiefstem Akzeptanzniveau gibt an, gar nicht zu reagieren, wie zur Auswahl steht.

Hypothese 7 wird für Szenario D bestätigt, da bei jeder der drei negativen Verhaltensabsichten über 64% der Probanden in den zwei tiefsten Akzeptanzkategorien liegen (vgl. Abb. A-7).

Abbildung A-7: Akzeptanz und negative Verhaltensabsichten (Szenario D)



Quelle: Eigene Darstellung

5. Effektgrösse und Teststärke

Da weder Mittelwerte noch t -, F - oder ChiQuadrat-Tests vorliegen, können die Effektgrösse und die Teststärke unter h7 nicht berechnet werden. Wie stark ein Effekt ist, kann aus den grafischen Trendanalysen abgelesen werden (vgl. Abb. A-6 und A-7).

A.6 Verzeichnis der Gesprächspartner

Name, Position des Gesprächspartners	Unternehmen / Institut	Datum, Ort	Zweck
Dr. Andreas Wittmer , Geschäftsführer	HSG Center for Aviation Competence	8. Mai 2007, St.Gallen	TG
Annemarie Zimmer , Manager Market Research	Swiss	10. Okt. 2007, Zürich-Kloten	QE
Barbara Josef , Head of Core Customer Management	Swiss	10. Mai 2006, Zürich-Flughafen	VQE
Bruno Walder , Sales Manager Switzerland	Singapore Airlines	30. März 2007, Zürich	QE
Christina Hollenweger , Marketing Manager	Singapore Airlines	30. März 2007, Zürich	QE
Christine Moch , Teamleader Advertising	Swiss	10. Okt. 2007, Zürich-Kloten	QE
Daniel Fanconi , Sales Executive	Singapore Airlines	30. März 2007, Zürich	QE
Edith Tieber , Head of Market Research	Unique	4. Mai 2007, Zürich-Flughafen	VE
Ivana Beslic , Assistant Passenger Services & Customer Relations	Swissport	23. Juli 2007, E-Mail Kontakt	VB
Jost Brändli , Manager Ground Services	Swiss	16. Juni 2006, telefonisch	VQE
Jürgen Imkamp , Rechtsanwalt mit Spezialgebiet Strafrecht	Baur + Imkamp Rechtsanwälte	15. Okt. 2006, telefonisch	TG
Kurt Breitenmoser , Psychiater, lic. oec.	Psychotherapie, St.Gallen	2. Okt. 2006, St.Gallen	TG
Dr. Luca Graf , Senior Manager Market Development Europe	Swiss	30. Apr. 2007, Zürich-Kloten	QE
Mathias Öchsli , Doktorand, lic. phil.	Lehrstuhl für Neuropsychologie, Universität Zürich	12. Okt. 2006, Zürich	TG
Nils Hartgen , Director / Head of E-Commerce	Swiss	30. Apr. 2007, Zürich-Kloten	QE
Reto Waltisberg , Teamleader Brand Management	Swiss	10. Okt. 2007, Zürich-Kloten	QE
Roberto Feijoo , Assistant Director Airline Handling Zurich & Complaint Coordinator	Jet Aviation	11. Apr. 2007, Zürich-Flughafen	QE

Legende: QE = Qualitative Empirie, TG = Theoretische Grundlagen, VB = Vorbereitung Beobachtung, VE = Vorbereitung Experiment, VQE = Vorbereitung qualitative Empirie

B Literaturverzeichnis

- Abernethy, A. M. und Butler, D. D. (1992). Advertising information: service versus products. *Journal of Retailing*, 68(4):398–419.
- Akers, R. L. (1977). *Deviant Behavior: A Social Learning Approach*. Wadsworth, Belmont CA, 2. Auflage.
- Anand, P., Holbrook, M. B. und Stephens, D. (1988). The formation of affective judgments: the cognitive-affective model versus the independence hypothesis. *Journal of Consumer Research*, 15(December):386–391.
- Anderson, J. C. und Gerbing, D. W. (1982). Some methods for respecifying measurement models to obtain unidimensional construct measurement. *Journal of Marketing Research*, 19(November):453–460.
- Anderson, P. F. (1983). Marketing, scientific progress, and scientific method. *Journal of Marketing*, 47(Fall):18–31.
- Andreassen, T. W. (1999). What drives customer loyalty with complaint resolution. *Journal of Service Research*, 1(4):324–332.
- Andreassen, T. W. (2000). Antecedents to satisfaction with service recovery. *European Journal of Marketing*, 34(1/2):156–75.
- Antonides, G. und van Raaij, W. F. (1998). *Consumer Behaviour: A European Perspective*. John Wiley & Sons, Chichester.
- Arndt, J. (1985). On making marketing science more scientific: role of orientations, paradigms, metaphors, and puzzle solving. *Journal of Marketing*, 49(3):11–23.
- Arnould, E. J. und Price, L. L. (1993). River magic: extraordinary experience and the extended service encounter. *Journal of Consumer Research*, 20(June):24–45.
- Athanassopoulos, A. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research*, 47(3):191–207.
- Atteslander, P. (2000). *Methoden der empirischen Sozialforschung*. De Gruyter, Berlin/New York.
- Babin, B. und Babin, L. (2001). Seeking something different? A model of schema typicality, consumer affect, purchase intentions and perceived shopping value. *Journal of Business Research*, 54(2):89–96.

- Bagozzi, P., Gopinath, M. und Nyer, P. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science*, 27(2):184–206.
- Baker, C. (2007). Checking in. *Airline Business*, (July):38–40.
- Ball, G. A., Trevino, L. K. und Sims, H. R., J. (1994). Just and unjust punishment: influences on subordinate performance and citizenship. *Academy of Management Journal*, 37(2):299–322.
- Bamberg, G., Baur, F. und Krapp, M. (2007). *Statistik*. Oldenbourg, München, 13. Auflage.
- Bamford, D. und Xystouri, T. (2005). A case study of service failure and recovery within an international airline. *Managing Service Quality*, 15(3):306–322.
- Bansal, H., Taylor, S. und St. James, Y. (2005). Migrating to new service providers: Toward a unifying framework of consumers' switching behaviors. *Journal of the Academy of Marketing Science*, 33(1):96–115.
- Barnes, J. W., Hadjimarcou, J. und Jacobs, R. S. (1999). Assessing the role of the customer in dyadic service encounters. *Journal of Customer Service in Marketing and Management*, 5(2):1–22.
- Baron, R. M. und Kenny, D. A. (1986). The mediator - moderator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51:1173–1182.
- Bearden, W. O. und Teel, J. E. (1983a). An investigation of personal influences on consumer complaining. *Journal of Marketing Research*, 20(February):21–28.
- Bearden, W. O. und Teel, J. E. (1983b). *Selected determinants of consumer satisfaction and complaint reports*. Association for Consumer Research, Provo, UT.
- Becker, G. (1968). Crime and punishment - an economic approach. *The Journal of Political Economy*, (76):169–217.
- Beckmann, J. (1992). *Handlungskontrolle und Leistung: Strategien und Mechanismen der Selbstregulation im Leistungshandeln*. Hogrefe, Göttingen.
- Bejou, D. und Palmer, A. (1998). Service failure and loyalty: an exploratory empirical study of airline customers. *Journal of Services Marketing*, 12(1):7–22.
- Belz, C. und Bieger, T. (2006). *Customer-Value. Kundenvorteile schaffen Unternehmensvorteile*. Thexis, Süddeutscher Verlag, Landsberg am Lech, 2. Auflage.
- Berger, R. und Wentzell, J. (2001). Optimizing online interaction: Cross-selling and upselling. *Customer Inter@action Solutions*, 19(10):44.

- Berry, L. L. (1980). Services marketing is different. *Business*, (May-June):24–29.
- Bieger, T. (2002). Geschäftsmodelle im Airline Bereich - Konvergenz oder Ausdifferenzierung. In *Jahrbuch der Schweizerischen Verkehrswirtschaft*. Universität St. Gallen (IDT-HSG), St. Gallen.
- Bieger, T. (2003). *Dienstleistungs-Management*. Haupt, Bern/Stuttgart/Wien.
- Bies, R. J. und Shapiro, D. L. (1987). Interactional fairness judgments: the influence of causal accounts. *Social Justice Research*, 1(2):199–218.
- Binggeli, U. und Pompeo, L. (2005). The battle for Europe's low-fare flyers [Elektronische Version]. *McKinsey Quarterly*, letzter Zugriff: 5. Januar 2008 von <http://www.slp.com/PDFs/Value.pdf>.
- Birnbaum, M. H. (1974). Reply to the devils advocates: Don't confound model testing and measurement. *Psychological Bulletin*, (81):854–859.
- Birnbaum, M. H. (2000). *Psychological Experiments on the Internet*. Academic Press, San Diego.
- Bitner, M., Booms, B. und Mohr, L. (1994). Critical service encounters: the employee's viewpoint. *Journal of Marketing*, 58(4):95–106.
- Bitner, M. J. (1990). Evaluating service encounters: the effects of physical surroundings and employee responses. *Journal of Marketing*, 54(2):69–82.
- Bitner, M. J. (1992). Servicescapes: the impact of physical surroundings on customers and employees. *Journal of Marketing*, 56(2):57–71.
- Bitner, M. J. (1995). Building service relationships: It's all about promises. *Journal of the Academy of Marketing Science*, 23(Fall):246–51.
- Bitner, M. J., Booms, B. H. und Tetreault, M. S. (1990). The service encounter: diagnosing favorable and unfavorable incidents. *Journal of Marketing*, 54(1):71–85.
- Blodgett, J. G., Granbois, D. H. und Walters, R. G. (1993). The effects of perceived justice on complaints' negative word-of-mouth behavior and repatronage intentions. *Journal of Retailing*, 399–427.
- Blumer, H. (1969). *Symbolic Interaction*. Prentice Hall, Englewood Cliffs.
- Bogdan, R. und Taylor, S. J. (1975). *Introduction to Qualitative Research Methods: A Phenomenological Approach to Social Sciences*. Wiley, New York.
- Boice, R. (1983). Observational skills. *Psychological Bulletin*, 93:3–29.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, New York.

- Bolton, R. N. und Drew, J. H. (1991). A multistage model of customers' assessments of service quality and value. *Journal of Consumer Research*, 17(4):875–884.
- Bortz, J. (2006). *Statistik für Sozialwissenschaftler*. Springer, Berlin, 6. Auflage.
- Boshoff, C. (1997). An experimental study of service recovery options. *International Journal of Service Industry Management*, 8(3):110–130.
- Boshoff, C. (1999). RECOVSAT: An instrument to measure satisfaction with transaction-specific service recovery. *Journal of Service Research*, 1(3):236–249.
- Bougie, R., Pieters, R. und Zeelenberg, M. (2003). Angry customers don't come back, they get back: the experience and behavioral implications of anger and dissatisfaction in services. *Journal of the Academy of Marketing Science*, 31(4):377–393.
- Boulding, W., Kalra, A., Staelin, R. und Zeithaml, V. A. (1993). A dynamic process model of service quality: From expectations to behavioral intentions. *Journal of Marketing Research*, 30(February):7–27.
- Bradley, G. V. H. (2003). Retribution: the central aim of punishment. *Journal of Law & Public Policy*, 27(1):19–31.
- Bredenkamp, J. (1980). *Theorie und Planung psychologischer Experimente*. Steinkopff, Darmstadt.
- Brodts, T. und Heé, J. (2005). *First world-wide Mobile Spam Study*. Universität St. Gallen (MCM), St. Gallen.
- Burns, A. C., Williams, L. A. und Maxham, J. (2000). Narrative text biases attending the critical incidents technique. *Qualitative Market Research: An International Journal*, 3(4):178–186.
- Buser, T. und Welte, B. (2006). *Customer Relationship Management für die Praxis*. Versus, Zürich.
- Butscher, S. (1998). *Handbuch Kundenbindungsmanagement*. IM Fachverlag, Ettlingen, D.
- Butterfield, K. D., Trevino, L. K. und Ball, G. A. (1996). Punishment from the manager's perspective: a grounded investigation and inductive model. *The Academy of Management Journal*, 39(6):1479–1512.
- Byrne, B. M. (2001). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Lawrence Erlbaum Associates, Mahwah.
- Campbell, D. T. und Stanley, J. C. (1966). *Experimental and Quasi-experimental Designs for Research*. Rand McNally, Chicago.
- Carlsmith, K. M., Darley, J. M. und Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, 83:284–299.

- Carlzon, J. (1987). *Moments of Truth*. Ballinger, Cambridge, MA.
- Carrier, M. (1995a). Realismus, wissenschaftlicher. In Blasche, S. und Mittelstrass, J., Hrsg., *Enzyklopädie Philosophie und Wissenschaftstheorie*, 506–507. Poeschel, Stuttgart.
- Carrier, M. (1995b). Relativismus. In Blasche, S. und Mittelstrass, J., Hrsg., *Enzyklopädie Philosophie und Wissenschaftstheorie*, 563–564. Poeschel, Stuttgart.
- Chang, T. Z. und Wildt, A. R. (1994). Price, product information, and purchase intention: an empirical study. *Journal of the Academy of Marketing Science*, 22(1):16–27.
- Chung, B. und Hoffman, K. D. (1998). Critical incidents: service failures that matter most. *Cornell Hotel and Restaurant Administration Quarterly*, 39(3):66–71.
- Chung-Herrera, B. G., Goldschmidt, N. und Hoffman, K. D. (2004). Customer and employee views of critical service incidents. *Journal of Services Marketing*, 18(4):241–254.
- Churchill, G. A. (1995). *Marketing Research: Methodological Foundations*. Dryden Press, Fort Worth, TX.
- Churchill, Gilbert A., J. und Surprenant, C. F. (1983). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research*, 19(November):491–504.
- Coby, P. (2007). The IT lever. *Airline Business*, (August):106.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum, Hillsdale.
- Coleman, J. S. (1982). *The Asymmetric Society*. Syracuse University Press, Syracuse.
- Coleman, J. S. (1990). *Foundations of Social Theory*. Belknap Press, Cambridge, Mass.
- Colgate, M. und Stewart, K. (1998). The challenge of relationships in services - a New Zealand study. *International Journal of Service Industry Management*, 9(5):454–468.
- Collis, R. (2007). Self-service options backed in air survey [Elektronische Version]. *International Herald Tribune*, letzter Zugriff: 14. Januar 2008 von <http://www.iht.com/articles/2007/12/13/travel/trfreq14.php>.
- Cook, T. D. und Campbell, D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Social Settings*. Rand McNally, Chicago.
- Cramer, E. und Kamps, U. (2007). *Grundlagen der Wahrscheinlichkeitsrechnung und Statistik*. Springer, Berlin.

- Cronin, J. J., Brady, M. K. und Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioural intentions in service environments. *Journal of Retailing*, 76(2):193–218.
- Cronin, J. J. und Taylor, J. R. (1992). Measuring service quality: a reexamination and extension. *Journal of Marketing*, 76(2):193–218.
- Cropanzano, R. und Folger, R. (1991). Procedural justice and worker motivation. In Steers, R. und Porter, L., Hrsg., *Motivation and Work Behavior*. McGraw-Hill, New York.
- Curren, M. T. und Folkes, V. S. (1987). Attributional influences on consumer's desires to communicate about products. *Psychology and Marketing*, 4(Spring):31–45.
- Czienskowski, U. (1996). *Wissenschaftliche Experimente: Planung, Auswertung, Interpretation*. Psychologie Verlags Union, Weinheim.
- Dabholkar, P. A. (1996). Consumer evaluation of new technology-based self-service options: an investigation of alternative models of service quality. *International Journal of Research in Marketing*, 13:234–246.
- Danaher, P. und Haddrell, V. (1996). A comparison of question scales used for measuring customer satisfaction. *International Journal of Service Industry Management*, 7(4):4–26.
- Day, R. L., Grabicke, K., Schaetzle, T. und Staubach, F. (1981). The hidden agenda of consumer complaining. *Journal of Retailing*, 57(Fall):86–106.
- de Coverly, E., Holme, N. O., Keller, A. G., Thompson, F. H. M. und Toyoki, S. (2002). Service recovery in the airline industry: is it as simple as 'failed, recovered, satisfied?'. *The Marketing Review*, 3(1):21–37.
- de Ruyter, K. und Wetzels, M. (2000). Customer equity considerations in service recovery: a cross-industry perspective. *International Journal of Service Industry Management*, 11(1):91–108.
- Deshpandé, R. (1983). Paradigms lost: on theory and method in research in marketing. *Journal of Marketing*, 47(4):101–110.
- Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as a basis for distributive justice? *Journal of Social Issues*, 31:137–149.
- Deutsch, M. und Krauss, R. M. (1965). *Theories in Social Psychology*. Basic Books, New York.
- Diehl, J. und Arbinger, R. (2001). *Einführung in die Inferenzstatistik*. Dietmar Klotz, Eschborn (Frankfurt), 3. Auflage.

- Diekmann, A. (1998). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Rowohlt, Reinbek.
- Dobruszkes, F. (2006). An analysis of European low-cost airlines and their networks. *Journal of Transport Geography*, 14(4):249–264.
- Doganis, R. (1995). The impact of liberalisation on european airline strategies and operations. *Journal of Air Transport Management*, 1:15.
- Doganis, R. (2001). *The Airline Business in the 21st Century*. Routledge, London.
- Doganis, R. (2002). *Flying Off Course: The Economics of International Airlines*. Routledge, London.
- Drenger, J. (2006). *Imagewirkungen von Eventmarketing*. Deutscher Universitäts-Verlag, Wiesbaden.
- Dubé, L. (1990). The power of affective reports in predicting satisfaction judgments. *Advances in Consumer Research*, 17:571–576.
- Dubé, L., Belanger, M.-C. und Trudeau, E. (1996). The role of emotions in health care satisfaction. *Journal of Health Care Marketing*, 16(2):45–51.
- Dubé, L. und Menon, K. (2000). Multiple roles of consumption emotions in post-purchase satisfaction with extended service transactions. *International Journal of Service Industry Management*, 11(3):287–304.
- Duller, C. (2007). *Einführung in die Statistik mit EXCEL und SPSS*. Physica-Lehrbuch, Heidelberg, 2. Auflage.
- Durkheim, E. (1895). *Die Regeln der soziologischen Methode*. Luchterhand, Neuwied, 4. (1976) Auflage.
- Dutka, A. F. (1994). *Handbook for Customer Satisfaction. A Complete Guide to Research, Planning & Implementation*. American Marketing Association, Lincolnwood.
- Easton, G. (2002). Marketing - a critical realist approach. *Journal of Business Research*, 55(2):103–109.
- Eccles, G. und Durand, P. (1998). Complaining customers, service recovery and continuous improvement. *Managing Service Quality*, 8(1):68–71.
- Edell, J. A. und Burke, M. C. (1987). The power of feelings in understanding advertising effects. *Journal of Consumer Research*, 14(December):421–433.

- Edvardsson, B. (1992). Service breakdowns: a study of critical incidents in an airline. *International Journal of Service Industry Management*, 3(4):17–29.
- Estelami, H. und Bergstein, H. (2006). The impact of market price volatility on consumer satisfaction with lowest-price refunds. *Journal of Services Marketing*, 20(3):169–177.
- Faul, F., Erdfelder, E., Lang, A.-D. und Buchner, A. (2007a). G*Power 3 - Version 3.0.6 für Mac OS X [Computer Programm]. *Heinrich Heine Universität Düsseldorf*, letzter Zugriff: 5. Januar 2008 von <http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3>.
- Faul, F., Erdfelder, E., Lang, A.-D. und Buchner, A. (2007b). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, (39):175–191.
- Fehr, E. (2005). Mit Neuroökonomik das menschliche Wesen ergründen. *Neue Zürcher Zeitung*, 146:29.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Row, Peterson, Evanston, IL.
- Feyerabend, P. K. (1975). *Wider den Methodenzwang. Skizze einer anarchistischen Erkenntnistheorie*. Suhrkamp, Frankfurt.
- Fischer, L. und Wiswede, G. (2002). *Grundlagen der Sozialpsychologie*. Oldenbourg Wissenschaftsverlag, München/Wien/Oldenbourg.
- Fisk, T. A., Brown, C. J., Cannizzaro, K. G. und Naftal, B. (1990). Creating patient satisfaction and loyalty. *Journal of Health Care Marketing*, 10(2):5–15.
- Fitzsimmons, J. A. und Fitzsimmons, M. J. (2001). *Service Management: Operations, Strategy, and Information Technology*. Irwin/McGraw-Hill, Boston.
- Flanagan, J. C. (1954). The Critical Incidents Technique. *Psychological Bulletin*, 51(4):327–359.
- Fleischer, J. (2004). CRM in Action. *Call Center Magazine*, 17(11):34.
- Folger, R. (1993). Reactions to mistreatment in work. In Murnighan, J. K., Hrsg., *Social Psychology in Organizations: Advances in Theory and Research*, 161–83. Prentice Hall, Engelwood Cliffs, NJ.
- Folger, R. und Cropanzano, R. (1998). *Organizational Justice and Human Resource Management*. Sage, Thousand Oaks, CA.
- Folkes, V. S. (1984). Consumer reactions to product failure: An attributional approach. *Journal of Consumer Research*, 10(March):398–409.

- Foster, S. L. und Cone, J. D. (1980). Current issues in direct observation. *Behavioral Assessment*, 2:313–338.
- Fournier, S. und Mick, D. G. (1999). Rediscovering satisfaction. *Journal of Marketing*, 63(4):5–23.
- Fram, E. und McCarthy, M. (1999). The true price of penalties. *Marketing Management*, 49–56.
- Fram, E. H. (1997). The customer penalty box. *Marketing Management*, 60–63.
- Fram, E. H. und Callahan, A. (2001). Do you know what the customer you penalized yesterday is doing today? A pilot analysis. *Journal of Services Marketing*, 15(6):496–509.
- Frick, A., Baechtiger, M. T. und Reips, U.-D. (2001). Financial incentives, personal information, and drop out in online studies. In Reips, U.-D. und Bosnjak, M., Hrsg., *Dimensions of Internet Science*. Pabst Science, Lengerich (D).
- Gabbott, M. und Hogg, G. (1994). Consumer behaviour and services: a review. *Journal of Marketing Management*, 10:311–324.
- Gephart, W. (1990). *Strafe und Verbrechen - Die Theorie Émile Durkheims*. Leske + Budrich, Opladen.
- Giese, J. und Cote, J. (2000). Defining consumer satisfaction. *Academy of Marketing Science Review (online)*.
- Gilbert, D. C. und Morris, L. (1995). The relative importance of hotels and airlines to the business traveller. *International Journal of Contemporary Hospitality Management*, 7(6):19–23.
- Gilovich, T., Keltner, D. und Nisbett, R. E. (2006). *Social Psychology*. W.W. Norton & Company Inc., New York.
- Girgensohn-Marchand, B. (1999). Ergebnisse der empirischen Kleingruppenforschung. In Schäfers, B., Hrsg., *Einführung in die Gruppensoziologie - Geschichte, Theorien, Analysen*. Quelle und Meyer, Heidelberg, Wiesbaden, 3. Auflage.
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday Anchor, New York.
- Goldstein, S. (2000). B’buster late fee settlement could hurt later. *Video Store*, 22(41):8.
- Goodwin, C. und Ross, I. (1992). Consumer responses to service failures: influence of procedural and interactional fairness perceptions. *Journal of Business Research*, 25:149–163.
- Göthlich, E. (2003). *Fallstudien als Forschungsmethode: Plädoyer für einen Methodenpluralismus in der deutschen betriebswirtschaftlichen Forschung*, Band 578. Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel.

- Greve, W. und Wentura, D. (1997). *Wissenschaftliche Beobachtung*. Psychologie Verlags Union, Weinheim, 2. Auflage.
- Grob, H. L. und Ahlert, D. (2006). *Internetökonomie: Ein interdisziplinärer Beitrag zur Erklärung und Gestaltung hybrider Systeme*. Vahlen, München.
- Grönroos, C. (1988). Service quality: the six criteria of good perceived service quality. *Review of Business*, 9(3):10–13.
- Grönroos, C. (2000). *Service Management and Strategy: Marketing the Moments of Truth in Service Competition*. Lexington Books, Lexington, MA, 2. Auflage.
- Grönroos, C. (2000). *Service Management and Marketing: A Customer Relationship Management Approach*. John Wiley and Sons, Crichester.
- Gröppel-Klein, A. (2001). Konsumentenverhaltensforschung. In Bruhn, M. und Homburg, C., Hrsg., *Gabler Marketing Lexikon*, 325–328. Gabler, Wiesbaden.
- Gross, P. (2005). *Die Multioptionsgesellschaft*. Suhrkamp, Frankfurt a. Main.
- Gukenbiehl, H. L. (1999). Formelle und informelle Gruppe als Grundformen sozialer Struktur- bildung. In Schäfers, B., Hrsg., *Einführung in die Gruppensoziologie - Geschichte, Theorien, Analysen*. Quelle und Meyer, Heidelberg, Wiesbaden, 3. Auflage.
- Güttler, P. O. (2003). *Sozialpsychologie: soziale Einstellungen, Vorurteile, Einstellungsänderungen*. Oldenbourg, München, 4. Auflage.
- Hamlin, R. P. (2005). The rise and fall of the Latin Square in marketing: a cautionary tale. *European Journal of Marketing*, 39(3/4):328–350.
- Harris, K. E., Mohr, L. A. und Kenneth, L. B. (2006). Online service failure, consumer attributions and expectations. *Journal of Services Marketing*, 20(7):453–458.
- Hart, C., Sasser, W. L., Earl, W., J. und Heskett, J. L. (1990). The profitable art of service recovery. *Harvard Business Review*, 68(July-August):148–156.
- Hart, H. L. A. (1982). Prolegomenon to the principles of punishment. In *Punishment and Responsibility*, Band 1. 5. Auflage.
- Havlena, W. J. und Holbrook, M. B. (1986). The varieties of consumption experience: comparing two typologies of emotion in consumer behavior. *Journal of Consumer Behavior*, 13:174–184.
- Hays, J. M. und Hill, A. V. (1999). The market share impact of service failures. *Production and Operations Management*, 8(3):208–220.

- Heckhausen, J. und Heckhausen, H. (2006). *Motivation und Handeln*. Springer, Berlin, 3. Auflage.
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. Wiley, New York.
- Hicks, C. R. (1993). *Fundamental Concepts in the Design of Experiments - Procedures for the Behavioral Sciences*. Oxford University Press, New York/Oxford.
- Hirsig, R. (2001). *Statistische Methoden in den Sozialwissenschaften*, Band 1. Seismo, Zürich.
- Hirsig, R. (2002). *Statistische Methoden in den Sozialwissenschaften*, Band 2. Seismo, Zürich.
- Hiskey, S. und Troop, N. A. (2002). Online longitudinal survey research: Viability and participation. *Social Science Computer Review*, 20(3):250–259.
- Hitzler, R. (1992). Der Goffmensch. Überlegungen zu einer dramatologischen Anthropologie. *Soziale Welt*, (43):449–461.
- Holbrook, M. B. und Batra, R. (1987). Assessing the role of emotions as mediators of consumer responses to advertising. *Journal of Consumer Research*, 14(December):404–420.
- Holloway, S. (2002). *Airlines: Managing to Make Money*. Ashgate, Aldershot, Burlington USA.
- Homans, G. C. (1961). *Social Behavior: Its Elementary Forms*. Social Forces. Harcourt, Brace & World, New York.
- Homans, G. C. (1972). *Grundfragen soziologischer Theorie*. Westdeutscher Verlag, Opladen.
- Homburg, C. (1989). *Exploratorische Ansätze der Kausalanalyse als Instrument der Marketingplanung*. Lang, Frankfurt.
- Homburg, C. und Giering, A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte. *Marketing ZFP*, 1:5–24.
- Honsell, H., Vogt, N. und Wiegand, W. (2003). *Basler Kommentar zum schweizerischen Privatrecht: Obligationenrecht*. Helbling und Lichtenhahn, Basel.
- Horovitz, B. (2001). Blockbuster to settle late-fee lawsuit agreement may be worth up to \$460 million to customers. *USA Today*, 01(June):B01.
- Huber, A. (1995). Beobachtung. In Roth, E., Hrsg., *Sozialwissenschaftliche Methoden*, Band 4. Oldenbourg, München.
- Huber, F., Vollhardt, K. und Heussler, T. (2006). *Emotion und Marke als Einflussfaktor der wahrgenommenen Preisfairness*. Wissenschaftliches Arbeitspapier, J. Gutenberg-Universität Mainz.
- Hui, M. und Bateson, J. E. G. (1991). Perceived control and the effects of crowding and consumer choice on the service experience. *Journal of Consumer Research*, 18(174-184).

- Hunt, S. D. (1983). General theories and the fundamental explananda of marketing. *Journal of Marketing*, 47(Fall):9–17.
- Hunt, S. D. (1991). Positivism and paradigm dominance in consumer research: toward critical pluralism and rapprochement. *Journal of Consumer Research*, 18(June):32–44.
- Hunt, S. D. (2003). *Controversy in Marketing Theory: For Reason, Realism, Truth and Objectivity*. Sharpe, Armonk N.Y.
- Hussy, W. und Jain, A. (2002). *Experimentelle Hypothesenprüfung in der Psychologie*. Hogrefe, Göttingen.
- Hwang, S. und Green, M. (1999). Restaurants get rude! *The Wall Street Journal*, 1, 14.
- IATA (2007). 2007 Corporate Air Travel Survey [Elektronische Version]. letzter Zugriff: 14. Januar 2008 von <http://www.iata.org/pressroom/briefings/2006-11-28-01.htm>.
- ICAO (1944). *Convention on International Civil Aviation*. International Civil Aviation Organization (ICAO). Doc. 2187, Chapter XXII, Article 96a. Montreal, December 7, 1944.
- Imhof, I. (2006). Meilen-Mathematik - Das Bonusprogramm Miles & More der Swiss verärgert Kunden. *Neue Zürcher Zeitung*, 273:71.
- Irlle, M. (1985). Konvergenz und Divergenz in Gruppen. In Frey, D. und Irlle, M., Hrsg., *Theorien der Sozialpsychologie*. Verlag Hans Huber, Bern.
- Jacobs, B. (2003). Versuchsplanung: Praktische Bedeutsamkeit der Effektstärke [Elektronische Version]. *Philosophische Fakultät der Universität des Saarlandes*, letzter Zugriff: 4. Januar 2008 von <http://www.phil.uni-sb.de/~jakobs/seminar/vpl/bedeutung/bedeut.htm>.
- James, L. R. und Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, 69:307–321.
- Janetzko, D., Hildebrandt, M. und Meyer, H. A. (2002). *Das Experimentalpsychologische Praktikum im Labor und WWW*. Hogrefe, Göttingen.
- Jäpel, W. (1985). *Die Qualität alternativer Rating-Formen und ihre Einflussgrößen*. Dissertation, Universität Regensburg.
- Jayawardhena, C., Wright, L. T. und Masterson, R. (2003). An investigation of online consumer purchasing. *Qualitative Market Research: An International Journal*, 6(1):58–65.
- Jensen, S. (1999). *Erkenntnis - Konstruktivismus - Systemtheorie*. Westdeutscher Verlag, Opladen, Wiesbaden.

- Johnson, R. und Wichern, D. (2002). *Applied Multivariate Statistical Analysis*. Prentice Hall, Englewood Cliffs.
- Johnson, R. L., Tsiros, M. und Lancioni, R. A. (1995). Measuring service quality: a systems approach. *Journal of Services Marketing*, 9(5):6–19.
- Johnston, R. und Mehra, S. (2002). Best practice complaint management. *Academy of Management Executive*, 16(4):145–154.
- Jones, T. O. und Sasser, W. E. (1995). Why satisfied consumers defect. *Harvard Business Review*, 73(6):88–102.
- Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, 51:296–312.
- Kahneman, D. und Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2):263–291.
- Kahneman, D. und Tversky, A. (2000). *Choices, Values, and Frames*. Cambridge University Press, New York.
- Kaiser, G. und Schöch, H. (2003). *Strafvollzug: Eine Einführung in die Grundlagen*. Springer, Heidelberg, 5. Auflage.
- Kant, I. (1797). *Metaphysik der Sitten*. In Weischedel, W. (1956), Hrsg., *Werke*, Band IV.
- Kaplan, S. E., Reckers, P. M. J. und Reynolds, K. D. (1986). An application of attribution and equity theories to tax evasion behavior. *Journal of Economic Psychology*, 7(4):461–476.
- Keates, N. (1998). Tricks of the frequent-traveling trade. *The Wall Street Journal*, 16(Section B):12.
- Keaveney, S. M. (1995). Customer switching behaviour in service industries: an exploratory study. *Journal of Marketing*, 59(April):71–82.
- Keller, K. L. (1993). Conceptualization, measuring, and managing customer-based brand equity. *Journal of Marketing*, 57(January):1–22.
- Kerlinger, F. N. (1973). *Foundations of Behavioural Research*. Holt, Rinehart & Winston, New York.
- Kim, Y. (2006). How attribution explains consumers' views of penalties. *The Journal of American Academy of Business*, 8(1):311–315.
- Kim, Y. K. (2003). *Customer evaluation of service organizations' penalties: an integrative framework based on cognitive, affective, and behavioral responses*. The George Washington University, Washington DC.

- Kim, Y. K. und Smith, A. K. (2005). Crime and punishment: examining customers' responses to service organizations' penalties. *Journal of Service Research*, 8(2):162–180.
- Kirk, R. (1995). *Experimental Design: Procedures for the Behavioral Sciences*. Brooks/Cole Publishing, Pacific Grove, CA, 3. Auflage.
- Klein, N. (1998). Bonussysteme für Kunden, Teil 2: Das Miles & More-Programm der Lufthansa. In Albers, S., Hassmann, S. und Tomczak, T., Hrsg., *Verkauf, Kundenmanagement, Vertriebssteuerung, E-Commerce*, 01.09. Gabler, Wiesbaden.
- Kley-Struller, A. (1993). *Wiedergutmachung im Strafrecht*, Band 1381. Europäische Hochschulschriften, Reihe II, Bern.
- Kotler, P. (2006). *Marketing Management*. Pearson Prentice Hall, Upper Saddle River, NJ, 12. Auflage.
- Kraemer, H. C. und Thiemann, S. (1987). *How many Subjects? Statistical Power Analysis in Research*. Sage, Newbury Park, CA.
- Kreuzpaintner, S. (2003). *Management unternehmensübergreifender Markensysteme innerhalb virtueller Dienstleistungsunternehmen am Beispiel Luftverkehr*. Dissertation Nr. 3194, Universität St. Gallen.
- Kroeber-Riel, W. und Weinberg, P. (2003). *Konsumentenverhalten*. Vahlen, München, 8. Auflage.
- Kromrey, H. (2006). *Empirische Sozialforschung: Modelle & Methoden der standardisierten Datenerhebung und Datenauswertung*. Lucius & Lucius, Stuttgart, 11. Auflage.
- Kuhl, J. (1987). Motivation und Handlungskontrolle: Ohne guten Willen geht es nicht. In Heckhausen, H., Gollwitzer, P. M. und Weinert, F. E., Hrsg., *Jenseits des Rubikon: Der Wille in den Humanwissenschaften*, 101–120. Springer, Berlin.
- Kumar, N., Scheer, L. K. und Steenkamp, J.-B. (1998). Interdependence, punitive capability, and the reciprocation of punitive actions in channel relationships. *Journal of Marketing Research*, 35(May):225–235.
- Kunz, K.-L. (2003). *Schweizerisches Strafrecht - Allgemeiner Teil*. Schulthess Juristische Medien, Zürich/Basel/Genf.
- Kuss, A. und Tomczak, T. (2004). *Käuferverhalten: Eine marketingorientierte Einführung*. Lucius & Lucius, Stuttgart, 3. Auflage.
- Laatz, W. (1993). *Empirische Methoden*. Verlag Harri Deutsch, Thun und Frankfurt.
- Lamnek, S. (2005). *Qualitative Sozialforschung*. Beltz, Weinheim, 4. Auflage.

- Lapidus, R. S. und Pinkerton, L. (1995). Customer complaint situations: an equity theory perspective. *Journal of Psychology and Marketing*, 12(2):105–119.
- Laucken, U., Schick, A. und Höge, H. (1996). *Einführung in das Studium der Psychologie. Eine Orientierungshilfe für Schüler und Studenten*. Klett-Cotta, Stuttgart, 7. Auflage.
- Lauer, T. (2004). *Bonusprogramme*. Springer, Berlin/Heidelberg/New York.
- LeBlanc, G. und Nguyen, N. (1996). Cues used by customers evaluating corporate image in service firms. *International Journal of Service*, 7(2):44–56.
- Lee, H., Lee, Y. und Yoo, D. (2000). The determinants of perceived service quality and its relationship with satisfaction. *Journal of Services Marketing*, 14(3):217–231.
- Leplin, J. (1986). Methodological realism and scientific rationality. *Philosophy of Science*, 53:31–51.
- Levesque, T. J. und McDougall, G. H. G. (2000). Service problems and recovery strategies: an experiment. *Canadian Journal of Administrative Sciences*, 17(1):20–38.
- Levy, J. (1996). Loss aversion, framing, and bargaining: the implications of prospect theory for international conflict. *International Political Science Review*, 17(2):179–195.
- Liljander, V. und Strandvik, T. (1997). Emotions in service satisfaction. *International Journal of Service Industry Management*, 8(2):148–169.
- Loehlin, J. C. (1987). *Latent Variable Models - An Introduction to Factor, Path, and Structural Analysis*. Lawrence Erlbaum Associates, Hillsdale.
- Loewenstein, G., Weber, E. U., Hsee, C. K. und Welch, N. (2001). Risk and feelings. *Psychological Bulletin*, 127(2):267–286.
- Love, R., Goth, J., Budde, F., Schiling, D. und Woffenden, B. (2006). Decline of the megahub. *Airline Business*, (October):72–76.
- Lovelock, C. H. (1992). *Managing Services - Marketing, Operations, and Human Resources*. Prentice Hall, Englewood Cliffs.
- Machleit, K. und Eroglu, S. (2000). Describing and measuring emotional response to shopping experience. *Journal of Business Research*, 49(2):101–111.
- Machleit, K. und Mantel, S. (2001). Emotional response and shopping satisfaction: moderating effects of shopper attributions. *Journal of Business Research*, 54(2):97–106.
- Malhotra, N. K. (2004). *Marketing Research - An Applied Orientation*. Prentice-Hall, Upper Saddle River, NJ.

- Manfreda, K. L. und Vehovar, V. (2002). Do mail and web surveys provide same results? *Advances in Methodology and Statistics*, 2(18):149–169.
- Mattila, A. S. und Patterson, P. G. (2004). Service recovery and fairness perceptions in collectivist and individualist contexts. *Journal of Service Research*, 6(4):336–346.
- Mattila, A. S. und Wirtz, J. (2004). Consumer complaining to firms: the determinants of channel choice. *Journal of Services Marketing*, 18(2):147–155.
- Mattila, A. S. und Wirtz, J. (2006). Arousal expectations and service evaluations. *International Journal of Service Industry Management*, 17(3):229–244.
- Maurer, P. (2003). *Luftverkehrsmanagement*. Oldenbourg, München, 3. Auflage.
- Maxham, J. G. und Netemeyer, R. G. (2002). Modeling customer perceptions of complaints handling over time: the effect of perceived justice on satisfaction and intent. *Journal of Retailing*, 78(4):239–252.
- Maxwell, S. E. und Delaney, H. D. (1990). *Designing experiments and analyzing data. A model comparison perspective*. Wadsworth, Belmont.
- Mayring, P. (2002). *Einführung in die qualitative Sozialforschung: eine Anleitung zu qualitativem Denken*. Psychologie Verlags Union, München, Weinheim.
- McCarthy, M. S. und Fram, E. H. (2000). An exploratory investigation of customer penalties: Assessment of efficacy, consequences, and fairness perceptions. *Journal of Services Marketing*, 14(6):479–501.
- McColl-Kennedy, J. R. und Sparks, B. A. (2003). Application of fairness theory to service failures and service recovery. *Journal of Service Research*, 5(3):251–266.
- McGuigan, F. J. (1993). *Experimental Psychology*. Prentice-Hall, Englewood Cliffs.
- McNaughton, R. B., Osborne, P. und Imrie, B. C. (2002). Market-oriented value creation in service firms. *European Journal of Marketing*, 39(9):990–1002.
- Meffert, H. (2000). *Marketing - Grundlagen Marktorientierter Unternehmensführung*. Gabler, Wiesbaden, 9. Auflage.
- Meffert, H. und Bruhn, M. (2000). *Dienstleistungsmarketing: Grundlagen - Konzepte - Methoden*. Gabler, Wiesbaden.
- Meier, R. F. und Johnson, W. T. (1977). Deterrence as social control: the legal and extralegal production of conformity. *American Sociological Review*, 42(April):292–304.

- Meleghy, T. (2001). Verhaltenstheoretische Soziologie: George Caspar Homans. In Morel, J., Bauer, E., Meleghy, T., Niedenzu, H.-J., Preglau, M. und Staubmann, H., Hrsg., *Soziologische Theorie - Abriss der Ansätze ihrer Hauptvertreter*. R. Oldenbourg Verlag, München, Wien, 7. Auflage.
- Merton, R. K. (1957). *Social Theory and Social Structure*. New York.
- Michel, S. (2001). Analyzing service failures and recoveries: a process approach. *International Journal of Service Industry Management*, 12(2):20 – 33.
- Mills, J. A. (2000). *Control: A History of Behavioral Psychology*. NY University Press, New York.
- Mittal, V., Ross, William T., J. und Baldasare, P. M. (1998). The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. *Journal of Marketing*, 62(January):33–47.
- Morris, M. W. und Moore, P. C. (2000). The lessons we (don't) learn: counterfactual thinking and organizational accountability after a close call. *Administrative Science Quarterly*, 45(4):737–765.
- Musch, J. und Reips, U.-D. (2000). A brief history of web experimenting. In Birnbaum, M. H., Hrsg., *Psychological Experiments on the Internet*, 61–85. Academic Press, San Diego.
- Myers, J. B., Pickersgill, A. D. und Van Metre, E. S. (2004). Steering customers to the right channels [Elektronische Version]. *The McKinsey Quarterly*, letzter Zugriff: 3. Januar 2008 von <http://www.mckinseyquarterly.com>.
- Nagel, T. (1991). *Die Grenzen der Objektivität*. Reclam, Stuttgart.
- Nolan, J., Ritchie, P. und Rowcroft, J. (2005). Small market air service and regional policy. *Journal of Transport Economics & Policy*, 39(3):363–378.
- Nolting, H.-P. und Paulus, P. (2005). *Psychologie lernen*. Beltz, Weinheim & Basel, 6. Auflage.
- Normann, R. (1991). *Service Management*. John Wiley & Sons, Chichester.
- O'Donohue, W. (1998). *Handbook of Behaviorism*. Academic Press, San Diego.
- Office of Fair Trading (1999). Trends in consumer complaints. Technical Report 47, Office of Fair Trading (OFT).
- Oliver, R. und Swan, J. E. (1989). Consumer perceptions of interpersonal equity and satisfaction in transactions: a field survey approach. *Journal of Marketing Research*, 53(April):24–35.
- Oliver, R. L. (1993). A conceptual model of service quality and service satisfaction: compatible goals, different concepts. *Advances in Services Marketing and Management*, 2:65–85.

- Oliver, R. L. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*. McGraw-Hill, New York.
- Oliver, R. L. (1999). Whence consumer loyalty? *Journal of Marketing*, 63(special issue):33–44.
- O'Malley, L. (1998). Can loyalty schemes really build loyalty? *Marketing Intelligence and Planning*, 16(1):47–55.
- O'Reilly, C. A. und Puffer, S. M. (1989). The impact of rewards and punishments in a social context: a laboratory and field experiment. *Journal of Occupational Psychology*, 62:41–53.
- o.V. (1995). Apostles and terrorists: a company's best friends and worst enemies [Elektronische Version]. *Harvard Business Review*, 73(6):96–98.
- o.V. (1999). Fares unfair. *The Economist*, 50.
- o.V. (2000). Überprüfung des Fahrverhaltens von Personenwagen in der Schweiz [Elektronische Version]. *BUWAL-Bericht 225*, letzter Zugriff: 3. Januar 2008 von <http://www.bafu.admin.ch>.
- o.V. (2001). Webbed Wings - Survey of Air Travel. *The Economist*, 21.
- o.V. (2002a). Bigger payouts for stranded passengers [Elektronische Version]. *BBC News*, letzter Zugriff: 4. Januar 2008 von <http://news.bbc.co.uk/2/hi/business/2356197.stm>.
- o.V. (2002b). The passengers are revolting [Elektronische Version]. *BBC News*, letzter Zugriff: 4. Januar 2008 von http://news.bbc.co.uk/2/hi/uk_news/2158520.stm.
- o.V. (2003). Kein Titel vorhanden. *The Washington Post*, E1–E2.
- o.V. (2004). BA sees profits more than double [Elektronische Version]. *BBC News*, letzter Zugriff: 5. Januar 2008 von <http://news.bbc.co.uk/1/hi/business/3991415.stm>.
- o.V. (2005a). Air passengers are left stranded [Elektronische Version]. *BBC News*, letzter Zugriff: 3. Januar 2008 von http://news.bbc.co.uk/2/hi/uk_news/scotland/4476568.stm.
- o.V. (2005b). Dreamweaver - Version 8.0 für Mac OS X [Computer Programm]. Von Macromedia Inc., letzter Zugriff: 3. Januar 2008 von <http://www.macromedia.com>.
- o.V. (2006a). *Brockhaus: Die Enzyklopädie in 30 Bänden*, Band 26. F.A. Brockhaus, Leipzig, Mannheim, 21. Auflage.
- o.V. (2006b). Marketingprofs [Elektronische Version]. Letzter Zugriff: 4. Januar 2008 von <http://search.marketingprofs.com>.
- o.V. (2006c). SPSS - Version 13.0.0 für Mac OS X [Computer Programm]. Von SPSS Inc., letzter Zugriff: 4. Januar 2008 von <http://www.spss.com>.

- o.V. (2006d). Zahlen und Fakten 2006. *Unique*.
- o.V. (2007a). European Online Travel Overview (3rd Edition) [Elektronische Version]. Von Phocus Wright Inc., letzter Zugriff: 4. Januar 2008 von <http://www.phocuswright.com>.
- o.V. (2007b). Schutz der Verbraucher vor irreführenden Flugangeboten im Internet: breit angelegte EU-Ermittlungen [Elektronische Version]. *EU-Verbraucherschutz, Referenz: IP/07/1694*, letzter Zugriff: 11. Januar 2008 von <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/07/1694&format=PDF&aged=0&language=DE&guiLanguage=de>.
- o.V. (2007c). Superlab [Web-basiertes Computer Programm]. Letzter Zugriff: 4. Januar 2008 von <http://www.superlab.com>.
- o.V. (2007d). WEXTOR [Web-basiertes Computer Programm]. Letzter Zugriff: 4. Januar 2008 von <http://psych-wextor.unizh.ch/wextor/en/login.php>.
- Parasuraman, A., Berry, L. L. und Zeithaml, V. A. (1991). Understanding customer expectations of service. *Sloan Management Review*, (Spring):9–48.
- Parson, T. und Sihls, E. A. (1962). Values, motives, and systems of action. In Parson, T. und Sihls, E. A., Hrsg., *Toward a General Theory of Action*, 2, 47 ff. Harvard University Press, Cambridge.
- Peeters, G. und Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. In Stroebe, W. und Hewstone, M., Hrsg., *European Review of Social Psychology*, Band 1, 33–60. Wiley, Chichester, UK.
- Peter, J. P. und Olson, J. C. (1983). Is science marketing? *Journal of Marketing*, 47(Fall):111–125.
- Polinsky, A. M. und Shavell, S. (1998). Punitive damages: An economic analysis. *Harvard Law Review*, 3(4):869–962.
- Pompl, W. (2002). *Luftverkehr - Eine ökonomische und politische Einführung*. Springer, Berlin, Heidelberg, New York, 4. Auflage.
- Popper, K. R. (1973). *Objektive Erkenntnis: Ein evolutionärer Entwurf*. Hoffmann und Campe, Hamburg.
- Preisendörfer, P. (2005). *Organisationssoziologie - Grundlagen, Theorien und Problemstellung*. Verlag für Sozialwissenschaften/GWV, Wiesbaden.
- Price, L., Arnould, E. J. und Deibler, S. L. (1995). Consumer's emotional responses to service encounters. The influence of the service provider. *International Journal of Service Industry Management*, 6(3):34–63.

- Rehberg, J. (2001). *Strafrecht II - Strafen und Massnahmen Jugendstrafrecht*. Schulthess Juristische Medien, Zürich, 7. Auflage.
- Rehberg, J. und Donatsch, A. (2001). *Verbrechenslehre*. Schulthess Juristische Medien, Zürich, 7. Auflage.
- Reichheld, F. F. (1996). Learning from customer defections. *Harvard Business Review*, 74(2):56.
- Reichheld, F. F. und Sasser, E. W. (1990). Zero defections: quality comes to services. *Harvard Business Review*, 68(September/October):105–111.
- Reips, U.-D. (2000). Das psychologische Experimentieren im Internet. In Batinic, B., Hrsg., *Internet für Psychologen*, Band 2. Hogrefe.
- Reips, U.-D. (2002). Internet-based psychological experimenting. *Social Science Computer Review*, 20(3):241–249.
- Reips, U.-D. und Neuhaus, P. (2002). WEXTOR: A Web-based tool for generating and visualizing experimental designs and procedures. *Behavior Research Methods*, 34(2):234–240.
- Reips, U.-D. und Stieger, S. (2004). Scientific LogAnalyzer: A Web-based tool for analyses of server log files in psychological research. *Behavior Research Methods*, 36(2):304–311.
- Rheinberg, F. (1995). *Motivation*. Kohlhammer, Stuttgart.
- Richins, M. (1997). Measuring emotions in the consumption experience. *Journal of Consumer Research*, 24(2):127–146.
- Richins, M. L. (1983). Negative word-of-mouth by dissatisfied consumers: a pilot study. *Journal of Marketing Research*, 47(Winter):68–78.
- Riemer, H. M. (1990). *Berner Kommentar: Band I, Einleitung und Personenrecht (die juristische Person)*. Stämpfli, Bern.
- Riklin, F. (2002). *Schweizerisches Strafrecht - Allgemeiner Teil I Verbrechenslehre*. Schulthess Juristische Medien, Zürich.
- Roberts, M. (1999). *A student's guide to Analysis of Variance*. Routledge, London.
- Roese, N. J. (1997). Counterfactual Thinking. *Psychological Bulletin*, 121(1):133–148.
- Ronan, W. W. und Latham, G. P. (1974). The reliability and validity of the critical incident technique: a closer look. *Studies in Personnel Psychology*, 6:53–64.
- Rosenberg, M. J. und Hovland, C. I. (1960). *Cognitive, Affective and Behavioral Components of Attitudes*. Yale University Press, New Haven, CT.

- Rössger, E. und Hünemann, K. (1965). *Einführung in die Luftverkehrspolitik*. Zürich.
- Roth, A. J. (1983). Robust trend tests derived and simulated: Analogs of the Welch and Brown-Forsythe tests. *Journal of the American Statistical Association*, 78(384):972–980.
- Rudolf-Sipötz, E. und Tomczak, T. (2001). *Kundenwert in Forschung und Praxis*. Thesis, St. Gallen.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178.
- Rust, R. T. und Oliver, R. L. (1994). *Service Quality: New Directions in Theory and Practice*. Thousand Oaks: Sage.
- Sandmo, A. (1993). Gary Becker's Contributions to Economics. *Scandinavian Journal of Economics*, 1(95):7–23.
- Sarris, V. (1999). *Einführung in die experimentelle Psychologie*. Pabst, Lengerich.
- Sarris, V. und Reiss, S. (2005). *Kurzer Leitfaden der Experimentalpsychologie*. Pearson Studium, München.
- Schäfers, B. (1999). Einführung. In Schäfers, B., Hrsg., *Einführung in die Gruppensoziologie - Geschichte, Theorien, Analysen*. Quelle und Meyer, Heidelberg, Wiesbaden, 3. Auflage.
- Schira, J. (2003). *Statistische Methoden der VWL und BWL - Theorie und Praxis*. Pearson Studium, München, 2. Auflage.
- Schmidt, I. D. (2004). *Kunden in Mehrkanalsystemen - Eine prozessorientierte Analyse des Kanalwahlverhaltens von Kunden in der Reisebranche*. Dissertation Nr. 2894, Universität St. Gallen.
- Schmidt, S. (1991). *Gedächtnis. Probleme und Perspektiven der interdisziplinären Gedächtnisforschung*. Suhrkamp, Frankfurt, M.
- Schneider, B. und Bowen, D. E. (1999). Understanding customer delight and outrage. *Sloan Management Review*, 41(1):35–45.
- Schnell, R., Hill, P. B. und Esser, E. (1999). *Methoden der empirischen Sozialforschung*. Oldenbourg, München, 6. Auflage.
- Schoefer, K. und Ennew, C. (2005). The impact of perceived justice on consumers' emotional responses to service complaint experiences. *Journal of Services Marketing*, 19(5):261–270.
- Schögel, M. (1998). Alternative Formen der Distribution: Electronic Commerce, Online Distribution und Direct Distribution. *Thesis-Fachzeitschrift für Marketing*, 1(88-92).

- Schögel, M. und Schulten, M. (2006). Wertorientierte Kundensteuerung in Mehrkanalsystemen. *Thesis-Fachzeitschrift für Marketing*, 4:37–42.
- Schönplflug, W. und Schönplflug, U. (1997). *Psychologie. Allgemeine Psychologie und ihre Verzweigungen in die Entwicklungs-, Persönlichkeits- und Sozialpsychologie*. Verlags Union, Weinheim, 4. Auflage.
- Schörcher, U. und Buchholz, R. (1994). Qualitätsmanagement in der Luftfahrt. In Massing, W., Hrsg., *Handbuch Qualitätsmanagement*. R. Oldenbourg Verlag, München/Wien.
- Seelmann, K. (2005). *Strafrecht - Allgemeiner Teil*. Helbling & Lichtenhahn, Basel/Genf/München, 2. Auflage.
- Shaw, S. (2002). *Airline Marketing and Management*. Ashgate, Hampshire (UK), 4. Auflage.
- Shostack, G. L. (1977). Breaking free from product marketing. *Journal of Marketing*, 41:73–80.
- Shostack, G. L. (1985). Planning the service encounter. In Czepiel, J. A., Solomon, M. R. und Surprenant, C. F., Hrsg., *The Service Encounter*, 243–254. Lexington Books, Lexington, MA.
- Shure, J. (1996). Beware those dangerous deadbeats that pay their bills. *Milwaukee Journal Sentinel*, 20(October), 2.
- Singh, J. (1990a). A typology of consumer dissatisfaction response styles. *Journal of Retailing*, 66(1):57–99.
- Singh, J. (1990b). Voice, exit, and negative word-of-mouth behaviors: an investigation across three service categories. *Journal of the Academy of Marketing Science*, 18(1):1–15.
- Singh, J. und Wilkes, R. E. (1996). When consumers complain: a path analysis of the key antecedents of complaint response estimates. *Journal of the Academy of Marketing Science*, 24:350–365.
- Skinner, B. F. (1938). *The Behavior of Organisms*. Appleton-Century, New York.
- Smith, A. K., Bolton, R. N. und Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of Marketing Research*, 36:356–372.
- Sobie, B. (2007). Keep on running. *Airline Business*, (8):58–85.
- Söderlund, M. (1998). Customer satisfaction and its consequences on customer behaviour revisited: The impact of different levels of satisfaction on word-of-mouth, feedback to the supplier and loyalty. *International Journal of Service Industry Management*, 9(2):169–188.
- Solomon, M. (2006). *Consumer Behaviour - Buying, Having and Being*. Prentice Hall, New Jersey, 7. Auflage.

- Sorell, T. (1994). The customer is not always right. *Journal of Business Ethics*, 13(11):913–918.
- Speer, T. L. (1996). They complain because they care. *American Demographics*, (May):13–14.
- Spreng, R. A. und Chiou, J. S. (2002). A cross-cultural assessment of the satisfaction formation process. *Journal of Consumer Research*, 36(7):829–839.
- Stafford, M. R. und Day, E. (1995). Retail service advertising: the effect of appeal, medium and service. *Journal of Advertising*, 24(1):58–71.
- Stahl, K.-H. (2000). Kundenloyalität kritisch betrachtet. In Hinterhuber, H. H. und Matzler, K., Hrsg., *Handbuch Kundenbindungsmanagement*, 319–335. Gabler, Wiesbaden.
- Stauss, B. (1998). Dienstleistungen als Markenartikel - etwas Besonderes? In Tomczak, T., Schögel, M. und Ludwig, E., Hrsg., *Markenmanagement für Dienstleistungen*. Thexis, St. Gallen.
- Stauss, B. (2004). Kundenfrustration - zur Marketingrelevanz der Frustrationstheorie. In Wiedmann, K.-P., Hrsg., *Fundierung des Marketing*, 63–86. Gabler, Wiesbaden.
- Stauss, B. und Neuhaus, P. (1997). The qualitative satisfaction model. *International Journal of Service Industry Management*, 8(3):236–249.
- Stauss, B., Schmidt, M. und Schoeler, A. (2005). Customer frustration in loyalty programs. *International Journal of Service Industry Management*, 16(3):229–252.
- Steiner, S. (2007). Belair und Edelweiss Air verweisen Dotcom-Flieger auf die hinteren Ränge. *Netzwoche*, 12(15-17).
- Stier, W. (1999). *Empirische Forschungsmethoden*. Springer, Berlin/Heidelberg/New York, 2. Auflage.
- Stratenwerth, G. (2005). *Schweizerisches Strafrecht, Allgemeiner Teil I: Die Straftat*. Stämpfli Verlag AG, Bern, 3. Auflage.
- Stratenwerth, G. und Kuhlen, L. (2004). *Strafrecht, Allgemeiner Teil I: Die Straftat*. Carl Heymanns Verlag, Köln, 5. Auflage.
- Sultan, F. und Simpson, M. C. (2000). International service variants: airline passenger expectations and perceptions of service quality. *Journal of Services Marketing*, 14(3):188–216.
- Surprenant, C. F. und Solomon, M. R. (1987). Predictability and personalization in the service encounter. *Journal of Marketing*, 51(April):73–80.
- Swanson, S. R. und Kelley, S. W. (2001). Service recovery attributions and word-of-mouth intentions. *European Journal of Marketing*, 35(1/2):194–211.

- Swedberg, R. (1990). *Economics and Sociology. Redefining Their Boundaries. Conversations with Economists and Sociologists*. Princeton, Oxford, U.K.
- Szymanski, D. und Henard, D. (2001). Customer satisfaction: a meta-analysis of the empirical evidence. *Journal of the Academy of Marketing Science*, 20(1):16–35.
- Taneja, N. K. (2002). *Driving Airline Business Strategies through Emerging Technology*. Ashgate, Aldershot, Burlington USA.
- Tax, S. S. und Brown, S. W. (1998). Recovering and learning from service failure. *Management Review*, 40(1):75–88.
- Tax, S. S., Brown, S. W. und Chandrashekar, M. (1998). Customer evaluations of service complaint experiences: Implications for relationship marketing. *Journal of Marketing*, 62(April):60–76.
- Taylor, S. (1994). Waiting for service: the relationship between delays and evaluations of service. *Journal of Marketing Research*, 58(April):56–69.
- Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin*, 110(1):67–85.
- Thaler, R., Tversky, A., Kahneman, D. und Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: an experimental test. *Quarterly Journal of Economics*, 112(2):647–661.
- Theil, M. (2002). *Versicherungsentscheidungen und prospect theory: die Risikoeinschätzung der Versicherungsnehmer als Entscheidungsgrundlage*. Springer, Wien.
- Theisen, M. R. (2005). *Wissenschaftliches Arbeiten*. Vahlen, München, 12. Auflage.
- Tomczak, T. (1992). Forschungsmethoden in der Marketingwissenschaft - Ein Plädoyer für den qualitativen Forschungsansatz. *Marketing ZFP*, 14(2):77–87.
- Tomczak, T. und Dittrich, S. (1997). *Erfolgreich Kunden binden*. Werd Verlag, Zürich.
- Tönnies, F. (1887). *Gemeinschaft und Gesellschaft: Abhandlung des Communismus und des Socialismus als empirischer Culturformen*. Reisland, Leipzig.
- Trechsel, S. und Noll, P. (2004). *Allgemeine Voraussetzungen der Strafbarkeit*. Zürich, 6. Auflage.
- Trommsdorff, V. (2003). *Konsumentenverhalten*. Kohlhammer, Stuttgart.
- Tversky, A. und Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211:453–458.

- Tversky, A. und Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business Research*, 59(October):251–78.
- Tversky, A. und Kahneman, D. (1991). Loss aversion in riskless choice: a reference-dependent model. *Quarterly Journal of Economics*, 106:1039–1061.
- Tyler, T. R. (1994). Psychological models of the justice motive: Antecedents of distributive and procedural justice. *Journal of Personality and Social Psychology*, 67:850–836.
- Ulrich, D. und Mayring, P. (1992). *Psychologie der Emotionen*. Kohlhammer, Stuttgart.
- Ulrich, H. (1981). Die Betriebswirtschaftslehre als anwendungsorientierte Sozialwissenschaft. In Geist, M. und Köhler, R., Hrsg., *Führung des Betriebs*. Schäffer-Poeschl, Stuttgart.
- Ulrich, H. (2001). Paradigmenwechsel in den Wirtschafts- und Sozialwissenschaften. In Stiftung zur Förderung der Systemorientierten Managementlehre St. Gallen, Schweiz, Hrsg., *Systemorientiertes Management*, 125–139. Haupt, Bern.
- Ulrich, P. und Hill, W. (1979). Wissenschaftstheoretische Grundlagen der Betriebswirtschaftslehre. In Raffée, H. und Abel, B., Hrsg., *Wissenschaftstheoretische Grundfragen der Wirtschaftswissenschaften*, 161–190. Vahlen, München.
- Utne, M. K. und Kidd, R. F. (1980). Equity and attribution. In Mikula, G., Hrsg., *Justice and Social Interaction: Experimental and Theoretical Contributions from Psychological Research*, 63–93. Springer-Verlag, New York.
- Van, J. (1999). AT&T rings up charges even if you don't call. *Rochester Democrat & Chronicle*, 1F – 6F.
- van Houten, R. (1983). Punishment: from the animal laboratory to the applied setting. In Axelrod, S. und Apsche, J., Hrsg., *The Effects of Punishment on Human Behavior*, 13–44. Academic Press, New York.
- Verma, H. V. (2001). Service failure and recovery in select industries. *Journal of Management Research*, 1(2).
- von Bartenwerffer, T. (2006). *Perceived Value in Open Service Environments*. Dissertation Nr. 3194, Universität St. Gallen.
- von Glasersfeld, E. (1985). Konstruktion der Wirklichkeit und des Begriffs der Objektivität. In Carl Friedrich von Siemens Stiftung, Hrsg., *Einführung in den Konstruktivismus*, 1–26. Oldenbourg, München.
- Walster, E., Berscheid, E. und Walster, G. W. (1973). New directions in equity research. *Journal of Personality and Social Psychology*, 25(Feb):151–176.

- Wardman, M. (1988). A comparison of revealed preferences and dated preferences models of travel behavior. *Journal of Transport Economics & Policy*, 22:71–91.
- Watson, D., Clark, L. A. und Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the panas scale. *Journal of Personality and Social Psychology*, 54:1063–1170.
- Weiner, B. (1985). Attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4):548–573.
- Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer Research*, 27(December):382–7.
- Weiner, B., Graham, S. und Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research*, 10:431–452.
- Weiser, C. R. (1995). Championing the customer. *Harvard Business Review*, 73(6):113–116.
- Weiss, H. M., Suckow, K. und Cropanzano, R. (1999). Effects of justice conditions on discrete emotions. *Journal of Applied Psychology*, 84(5):786–794.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik. Ein Lehrbuch zur Psychologischen Methodenlehre*. Hogrefe, Göttingen.
- White, C. und Yu, Y.-T. (2005). Satisfaction emotions and consumer behavioral intentions. *Journal of Services Marketing*, 19(6):411–420.
- Winer, B. J., Brown, D. R. und Michels, K. M. (1991). *Statistical principles in experimental design*. McGraw-Hill, Boston.
- Winkler, A. C. und McCuen, J. R. (1999). *Writing the Research Paper*. Harcourt College Publishers, Forth Worth, TX, 5. Auflage.
- Wirtz, J. und Bateson, J. E. G. (1999). Consumer satisfaction with services: integrating the environment perspective in services marketing into the traditional disconfirmation paradigm. *Journal of Business Research*, 44(1):55–66.
- Wirtz, J. und Chew, P. (2002). The effects of incentives, deal proneness, satisfaction and tie strength on word-of-mouth behaviour. *International Journal of Service Industry Management*, 13(2):141–162.
- Wirtz, J. und Mattila, A. S. (2001). The impact of expected variance in performance on satisfaction process. *International Journal of Service Industry Management*, 12(4):342–358.

- Wirtz, J. und Mattila, A. S. (2004). Consumer responses to compensation, speed of recovery and apology after a service failure. *International Journal of Service Industry Management*, 15(2):150–166.
- Witte, E. (1977). Lehrgeld für empirische Forschung: Notizen während einer Diskussion. In Köhler, R., Hrsg., *Empirische und handlungstheoretische Forschungskonzeptionen in der Betriebswirtschaftslehre*, 269–281. Schäffer-Poeschel, Stuttgart.
- Wong, A. und Sohal, A. (2003). A critical incident approach to the examination of customer relationship management in a retail chain: an exploratory study. *Qualitative Market Research: An International Journal*, 6(4):248–262.
- Woodyard, C. (1997). Circuit city tacks 15% fee on some returns. *USA Today*, 1.
- Yin, R. K. (1994). *Case Study Research - Design and Methods*. Sage Publications, London/New Delhi.
- Ying, H. L. und Chung, C. M. (2007). The effects of single-message single-source mixed word-of-mouth on product attitude and purchase intention. *Asia Pacific Journal of Marketing*, 19(1):75–86.
- Zajonc, R. B. (1965). Social facilitation. *Science*, 149(3681):269–274.
- Zajonc, R. B. (1985). Emotion and facial efference. *Science*, 228:15–21.
- Zeithaml, V. (2000). Service quality, profitability and the economic worth of customers: what we know and what we need to learn. *Journal of the Academy of Marketing Science*, 28(1):67–85.
- Zeithaml, V. A., Berry, L. L. und Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, 60(April):31–46.
- Zeithaml, V. A. und Bitner, M. J. (2000). *Service Marketing*. Irwin/McGraw Hill, New York.

Lebenslauf

János Heé, geboren am 8. August 1977 in Zürich

Ausbildung

- 10/04 – 09/08 Universität St. Gallen (HSG)
Doktorandenstudium, Fachprogramm Marketing
- 10/98 – 03/04 Universität St. Gallen (HSG)
Grund- und Hauptstudium, Fachrichtung Betriebswirtschaftslehre:
Informations- und Technologiemanagement (ITM)
- 08/93 – 01/98 Kantonsschule Hottingen, Zürich
Wirtschaftsmatura (Typus E)

Berufserfahrung

- 04/04 – 06/07 Intrado Communications, Zug / Colorado
Business Development Manager
- 07/03 – 10/03 ExecuJet Aviation Group, Zürich-Flughafen
Praktikum im Bereich Business Aviation
- 04/01 – 07/03 Institut für Wirtschaftsinformatik, Universität St. Gallen
Hilfsassistent im Bereich Webservices (Teilzeit)
- 10/00 – 03/01 Arthur D. Little Consulting, Zürich
Praktikum im Bereich M-Commerce

