



Bundesministerium
für Verkehr, Bau
und Stadtentwicklung



Leitfaden für die Durchführung von Kundenzufriedenheitsbefragungen im ÖPNV



Auftraggeber:
Bundesministerium für Verkehr,
Bau und Stadtentwicklung
Referat UI 31
Invalidenstraße 44
10115 Berlin
Internet: www.bmvbs.de

Projektträger:
TÜV Rheinland Consulting GmbH
Forschungsmanagement BMVBS
Am Grauen Stein 33
51105 Köln
Internet: www.tuv.com

Beratergremium:
Verband Deutscher Verkehrsunternehmen e. V. (VDV)
VDV-Arbeitssteam „Kundenzufriedenheit“
Kamekestraße 37-39
50672 Köln
Internet: www.vdv.de

Inhalt/Redaktion:
Dipl.-Stat. Marcus Bäumer
Dipl.-Soz. Manfred Pfeiffer
IVT Research GmbH
Quadrat M 4, 10
68161 Mannheim
Internet: www.ivt-verkehrsforschung.de

Inhaltsverzeichnis

Aktuelle Erkenntnisse aus dem Forschungsprojekt		
	„Bedeutung von Kundenzufriedenheit im ÖPNV“	4
1	Einführung	8
	1.1 Zum Anliegen des Praxisleitfadens	8
	1.1.1 Grundlage und Zielsetzung des Leitfadens.....	8
	1.1.2 Zur Bedeutung der Kundenzufriedenheit	8
	1.2 Das Konzept der Kundenzufriedenheit	9
	1.3 Zusammenhang von Kundenzufriedenheit und wirtschaftlichem Erfolg.....	11
2	Überblick über die Methoden zur Kundenzufriedenheits- und Qualitätsmessung	13
	2.1 Beobachtung.....	14
	2.2 Inhaltsanalyse	16
	2.3 Kundenzufriedenheitsmessung mittels Befragung: Ereignisorientierte Verfahren	17
	2.3.1 Sequentielle Ereignismethode.....	18
	2.3.2 Critical Incident-Technik	20
	2.3.3 Frequenz-Relevanz-Analyse von Problemen (FRAP).....	22
	2.4 Kundenzufriedenheitsmessung mittels Befragung: Merkmalsorientierte Verfahren	23
	2.4.1 Multiattributive Ansätze	23
	2.4.2 Penalty-Reward-Faktoren-Ansatz	25
	2.4.3 Dekompositionelle Verfahren	27
	2.5 Kundenzufriedenheitsmessung mittels Befragung: Gruppendiskussionen	29
	2.6 Analyse prozessproduzierter Daten.....	30
3	Durchführung von Kundenzufriedenheitserhebungen	32
	3.1 Methodische Hinweise zur Durchführung von Kundenzufriedenheitsbefragungen.....	32
	3.1.1 Festlegung der Untersuchungsziele.....	32
	3.1.2 Befragungsdesign.....	32
	3.1.3 Vorbereitung und Durchführung der Befragung	39
	3.1.4 Datenaufbereitung und Auswertung.....	42

3.2	Beschreibung der zugrunde liegenden Erhebung	43
3.3	Multiattributive Verfahren	44
3.3.1	Einstellungs- vs. zufriedenheitsorientierte Messung	44
3.3.2	Direkte vs. indirekte Messung	49
3.3.3	Ein- vs. Zweikomponentenansatz	50
3.3.4	Penalty-Reward-Faktoren-Analyse	53
3.3.5	Messung der Kundenbindung.....	56
3.4	Critical Incident-Verfahren	57
3.4.1	Erhebung	57
3.4.2	Auswertung.....	58
3.4.3	Ergebnisse.....	59
4	Bewertung der Verfahren hinsichtlich der Möglichkeiten zur Ableitung von Maßnahmen	62
5	Praxistipps	68
	Literatur	72

Aktuelle Erkenntnisse aus dem Forschungsprojekt „Bedeutung von Kundenzufriedenheit im ÖPNV“

Im Bereich der ÖPNV-Praxis gibt es eine Reihe von Fragestellungen in Verbindung mit Studien zur Kundenzufriedenheit, die bislang nicht eindeutig gelöst sind. Teilweise kursieren Vermutungen und Hypothesen, die durch einzelne Studien gestützt und durch andere widerlegt werden. Daraus resultiert die Frage nach der Validität von Untersuchungsergebnissen zur Kundenzufriedenheit und deren Eignung zur Ableitung von Maßnahmen der Angebotsplanung. Sind hohe Kundenzufriedenheitswerte zugleich ein Synonym für einen attraktiven ÖPNV und damit als Instrument zur Nachfragesteigerung geeignet? Bislang fehlt ein valider Nachweis der Hypothese, wonach eine Steigerung der Kundenzufriedenheit einen Nachfragezuwachs bewirkt. Somit ergeben sich zwei Kernfragen, denen in einem Forschungsprojekt des BMVBS nachgegangen wurde:

- Wie wird die Qualität des ÖPNV wahrgenommen und angenommen, und wie werden bestimmte angebotsverändernde Maßnahmen über alle Kunden hinweg beurteilt bzw. wie wird der ÖPNV vom Rest der Bevölkerung, also potenziellen Neukunden gesehen, die manchmal nur sporadisch oder gar nicht mit dem ÖPNV in Berührung kommen, sich aber dennoch eine Meinung bilden und diese auch kundtun?

und daraus resultierend die für jedes Unternehmen entscheidende Frage:

- Führt höhere Kundenzufriedenheit tatsächlich auch zu einem höheren Unternehmenserfolg?

Das Bundesministerium für Verkehr, Bau und Stadtentwicklung (BMVBS) hat zu diesem Zweck ein Forschungsvorhaben in Auftrag gegeben, welches auf Anregung und in Zusammenarbeit mit dem VDV durchgeführt wurde und zum Ziel hatte, die Methodik und Anwendungsmöglichkeiten von Kundenzufriedenheitserhebungen systematisch zu analysieren. Um sowohl die erhebungsmethodischen als auch die inhaltlichen Analysen auf empirisch belastbare Daten zu gründen, wurde im Rahmen des Projektes eine standardisierte Kundenzufriedenheitsbefragung mit insgesamt 4.855 Interviews durchgeführt.

Das besondere am gewählten Untersuchungsdesign war, dass

- i) die Untersuchung als Panel angelegt war, d.h. es konnten im Abstand von einem Jahr dieselben Personen erneut zu den gleichen Themen befragt werden

- ii) in die Erhebung alle drei relevanten „Nutzergruppen“ des ÖPNV (Vielnutzer, Gelegenheitsnutzer und Nichtnutzer) einbezogen wurden
- iii) die Erhebung in drei Städten stattfand, in denen zwischen den beiden Erhebungswellen angebotsverändernde Maßnahmen im ÖPNV eingeführt wurden.

So konnte neben der Analyse geeigneter Methoden auch festgestellt werden, welche spezifische Wirkung von den jeweiligen Maßnahmen in den Untersuchungsräumen auf die Kundenzufriedenheit ausgingen. Darüber hinaus wurde der Zusammenhang zwischen Kundenzufriedenheit und Unternehmenserfolg auf der Ebene der Verkehrsunternehmen bzw. -verbünde analysiert. Hierzu wurden Ergebnisse des ÖPNV-Kundenbarometers mit bestimmten - vom VDV jährlich erhobenen - wirtschaftlichen Unternehmenskennziffern (z.B. Unternehmensertrag und -fahrten) verknüpft.

Die Ergebnisse der Studie sind im Schlussbericht für das BMVBS dokumentiert (FOPS-Projekt FE 70.796/2006). An dieser Stelle sollen zusammenfassend einige ausgewählte Erkenntnisse vorgestellt werden, die für die Arbeit mit Kundenzufriedenheitsuntersuchungen hilfreich erscheinen.

- Im Hinblick auf den Unternehmensertrag und die Anzahl beförderter Fahrgäste (Unternehmensfahrten) konnte im Rahmen der Analysen erstmals ein signifikanter Zusammenhang mit der Kundenzufriedenheit nachgewiesen werden.
- Die Paneluntersuchungen bestätigen den Wirkungszusammenhang zwischen der Kundenzufriedenheit und Nutzungsintensität. Es konnte gezeigt werden, dass Personen, die den ÖPNV zu einem bestimmten Zeitpunkt unterdurchschnittlich bewerten, ihn zukünftig eher seltener nutzen. Ebenso bestätigt sich der umgekehrte Fall, dass bei einer überdurchschnittlichen Bewertung eine zukünftige stärkere Nutzung wahrscheinlich ist.
- Die Untersuchungen haben weiterhin gezeigt, dass keine pauschal übertragbaren Aussagen getroffen werden können, in welchem Maße sich eine angebotsverändernde Maßnahme auf die Kundenzufriedenheit auswirkt, da dies in hohem Maße durch gebietsspezifische Rahmenbedingungen beeinflusst wird.
- Die Kundenzufriedenheit wird durch das lokal vorhandene Anspruchsniveau mitbestimmt. Da mit einem verbesserten Angebot auch die Ansprüche der Kunden im Zeitverlauf steigen, muss das Thema Steigerung oder zumindest Aufrechterhaltung der Kundenzufriedenheit auch als ein zeitabhängiger Prozess gesehen werden.

- Bestimmte angebotsverändernde Maßnahmen wie z.B. die Einführung eines dynamischen Fahrgastinformationssystems werden nicht nur von den Kunden des ÖPNV wahrgenommen werden, sondern auch von deutlich über 50 % der Nicht-Nutzer. Ihre Einführung kann damit auch eine generelle Image fördernde Wirkung besitzen.
- Auf die direkte Frage nach den Ursachen für eine gesteigerte ÖPNV-Nutzung werden neben angebotsverbessernden Maßnahmen auch externe Rahmenbedingungen wie schlechte Parkplatzsituation oder steigende Benzinpreise, und insbesondere Veränderungen der persönlichen Lebensumstände (Auto an-, abgeschafft, Umzug, Arbeitsplatzwechsel) angeführt.
- In den statistischen Analysen konnte ein signifikanter Zusammenhang zwischen der Globalzufriedenheit und positiven bzw. negativen Erlebnissen im ÖPNV identifiziert werden.
- Die Analysen zeigten zudem, dass negative Ereignisse etwas häufiger weitererzählt werden als positive. Bei den negativen Erlebnissen entfallen ca. 45 % der Nennungen auf Kontakte bzw. Konflikte mit anderen Personen (Personal oder andere Fahrgäste). Die Beschwerdequote bei Personen mit negativen Erlebnissen beträgt etwa 29 %.
- Neuere Ansätze wie z.B. die Critical-Incident-Technik sind geeignet, um solche Schwachstellen bzw. Ansatzpunkte für Verbesserungen zu identifizieren. Entsprechende Maßnahmen im Bereich Personal/Kundenumgang können zumeist schnell und teilweise mit geringem Aufwand umgesetzt werden. Werden diese Verbesserungen noch von entsprechenden Kommunikationsmaßnahmen flankiert, so kann (auch durch die einsetzende Mund-zu-Mund-Propaganda) eine Verbesserung des Images eines Unternehmens erreicht werden.
- Ein Vergleich der Einschätzung von ÖPNV-Experten auf der einen und den Befragten aus der Kundenzufriedenheitserhebung auf der anderen Seite zeigte neben vielen Übereinstimmungen auch gravierende Unterschiede. So werden in beiden Gruppen die „Pünktlichkeit“ und die „Sicherheit der Anschlüsse“ als zentrale Einflussgrößen auf die Kundenzufriedenheit gesehen. Andererseits wird z.B. die persönliche Sicherheit an Haltestellen von der Bevölkerung als deutlich wichtiger im Hinblick auf die Kundenzufriedenheit eingestuft als von den Vertretern der Verkehrsunternehmen und Verbände. Hier zeigt sich die Relevanz von unabhängig durchgeführten repräsentativen Kundenbefragungen.
- Die Analysen haben allerdings auch gezeigt, dass häufig externe Faktoren (wie Übergriffe im Bereich der U-Bahn bzw. die flächendeckende Einführung von Kurzarbeit), die gar nicht oder nur in geringem Maße vom Verkehrsunternehmen oder Verbund beeinflusst werden können, einen Einfluss auf die Kundenzufriedenheit bzw. die ÖPNV-Nutzung haben können.

Deshalb bietet es sich an, Längsschnitterhebungen turnusmäßig (jedes Quartal oder Jahr) durchzuführen oder zu bestimmten inhaltlich begründeten Zeitpunkten, z.B. wenn sich grundlegende Angebotsparameter geändert haben. Der Einfluss externer temporärer Rahmenbedingungen (z.B. Großbaustelle oder gewaltsamer Übergriff), kann bei Längsschnitterhebungen besser identifiziert und quantifiziert werden als bei einmalig durchgeführten Erhebungen. Das Projekt hat gezeigt, dass zum Thema „Einfluss von externen Faktoren auf die Kundenzufriedenheit“ und „Änderung der Lebensumstände und Auswirkungen auf das Mobilitätsverhalten“ noch weiterer Forschungsbedarf besteht.

Als Arbeitshilfe zur adäquaten Planung, Vorbereitung und Durchführung von Kundenzufriedenheitsuntersuchungen werden im vorliegenden Papier Erkenntnisse des Forschungsvorhabens praxisbezogen zusammengefasst und Empfehlungen für die Eignung der unterschiedlichen Verfahren für Anwendungen im Bereich des ÖPNV abgeleitet. Damit wird die Zielsetzung verfolgt, in kompakter Form einen Überblick über die Methoden und Einsatzmöglichkeiten der Kundenzufriedenheitsmessung zu vermitteln und den Praktikern in den Unternehmen Anregungen bei der Auswahl der Verfahren zu geben. Anhand von Beispielen werden die unterschiedlichen Anwendungen illustriert und im letzten Kapitel Praxistipps u.a. für die Vergabe und Durchführung von Kundenzufriedenheitserhebungen an externe Büros gegeben.

1 Einführung

1.1 Zum Anliegen des Praxisleitfadens

1.1.1 Grundlage und Zielsetzung des Leitfadens

Zur Messung der Zufriedenheit von Kunden mit dem Leistungsangebot steht mittlerweile eine breite Palette von methodischen Ansätzen zur Verfügung. Das Ziel des vorliegenden Leitfadens besteht zum einen darin, die verschiedenen Methoden der Kundenzufriedenheitsmessung überblicksartig darzustellen und miteinander zu vergleichen. Zum anderen sollen insbesondere die auf Befragungen von Kunden basierenden Verfahren detailliert beschrieben und allgemeine methodische Hinweise zur Durchführung solcher Befragungen gegeben werden. Mit diesem Leitfaden soll der Leser somit in die Lage versetzt werden, die nachfolgenden Fragen beantworten zu können:

- Welche Erhebungsmethoden für die Kundenzufriedenheit gibt es und wie funktionieren sie?
- Für welchen Zweck ist welche Methode geeignet? Was sind die Vor- und die Nachteile?
- Wie kann ich aus den Ergebnissen von Kundenzufriedenheitserhebungen geeignete Maßnahmen ableiten?

Dabei soll dieser Leitfaden nicht den Anspruch erheben, eine umfassende Darstellung aller vorgestellten Verfahren zu leisten, dazu dienen die weiterführenden Literaturhinweise. Eine vollständige Dokumentation des Forschungsvorhabens liegt im Schlussbericht für das BMVBS vor (FOPS-Projekt FE 70.796/2006) und kann über das IVT bezogen werden.

1.1.2 Zur Bedeutung der Kundenzufriedenheit

Untersuchungen zur Kundenzufriedenheit werden wie in vielen anderen Dienstleistungsbereichen auch im ÖPNV-Markt seit Jahren angewandt. Sie werden sowohl auf Unternehmensebene als auch im regionalen Bezug z.B. durch Verbundorganisationen durchgeführt. Daneben gibt es überregionale Studien, die auch Vergleiche auf nationaler oder internationaler Ebene ermöglichen. Trotz der vielfältigen Anwendungen ist bisher keine empirisch belegte Studie bekannt, in der eine ökonomische Nutzenstiftung einer steigenden Kundenzufriedenheit nachgewiesen wurde. Wenngleich die Vermutung nahe liegt, dass eine hohe Kundenzufriedenheit mit dem wirtschaftlichen Erfolg eines Unternehmens korreliert, ist diese Annahme also bisher nicht be-

legt. Ebenso besteht eine zentrale Fragestellung dahingehend, ob und auf welche Weise aus Studien zur Kundenzufriedenheit konkrete Handlungsmaßnahmen zur Verbesserung der Angebotsqualität mit dem Ziel der Kundenbindung und Kundengewinnung abgeleitet werden können.

Gerade vor dem Hintergrund begrenzter Finanzmittel für den ÖPNV soll das Dienstleistungsangebot des ÖPNV so attraktiv sein, dass es auch wahlfreie Verkehrsteilnehmer anspricht. Um dieses Ziel zu erreichen, sind Verkehrsunternehmen, -verbände und Aufgabenträger bestrebt, mit dem Verkehrsangebot eine bestmögliche Nachfragewirkung zu erzielen. Insofern stellt sich für alle ÖPNV-Partner die Frage, ob und inwieweit Maßnahmen zur Erhöhung der Kundenzufriedenheit zugleich dazu dienen, die Einnahmesituation bzw. Wirtschaftlichkeit zu verbessern.

Bei der Vergabe von Verkehrsleistungen durch Aufgabenträger im Wettbewerb werden in der Regel neben den Kosten kundenorientierte Qualitätskriterien einbezogen, um zu vermeiden, dass durch Billigangebote mit unzureichender Qualität Kundenpotenziale verloren gehen oder nicht ausgeschöpft werden. Auch in diesem Kontext kommen daher Kundenzufriedenheitsuntersuchungen als Instrumente zur Qualitätssteuerung zum Einsatz.

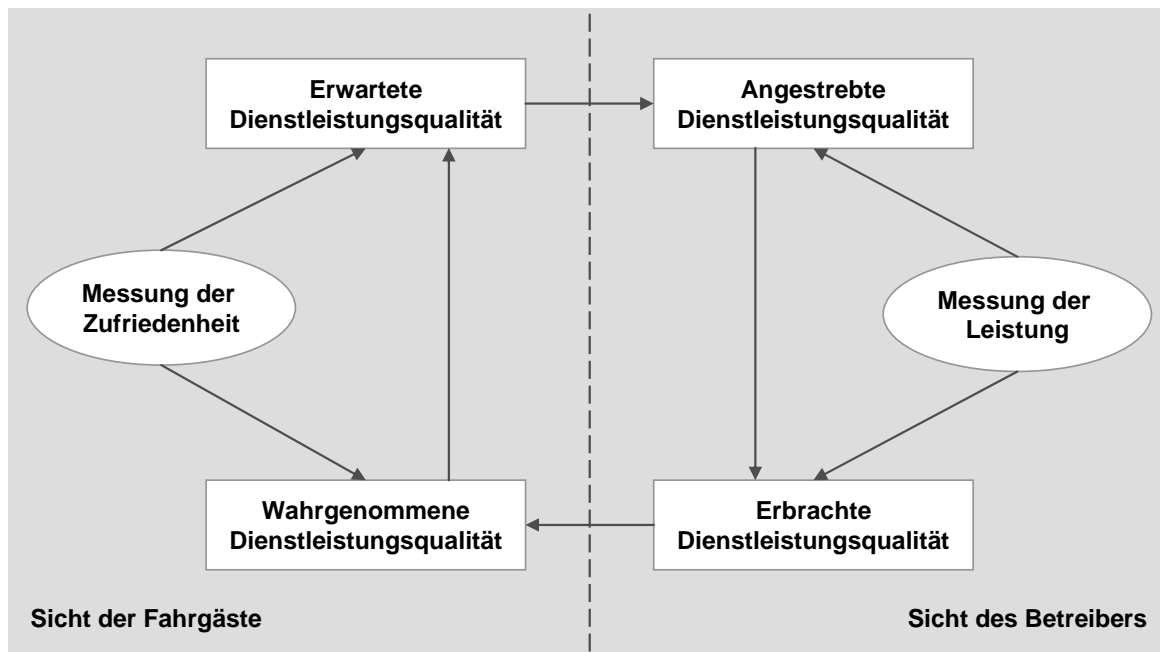
1.2 Das Konzept der Kundenzufriedenheit

Ganz allgemein wird unter Kundenzufriedenheit das Ergebnis eines individuellen Vergleichsprozesses verstanden, bei dem die wahrgenommene Erfahrung nach Gebrauch eines Produktes oder einer Dienstleistung mit den Erwartungen oder Wünschen vor der Nutzung verglichen wird. Entspricht die wahrgenommene Leistungsqualität der Erwartung oder übertrifft diese, entsteht Zufriedenheit. Bei der Kundenzufriedenheit handelt es sich nach gängiger Auffassung um ein „Nachkaufphänomen“, um das Ergebnis einer Ex-post-Beurteilung, die ein konkretes, selbst erfahrenes Konsumerlebnis voraussetzt (Stauss & Seidel 2006).

Diese Definition der Kundenzufriedenheit als Reaktion auf einen Ist-Soll-Vergleich entspricht dem Grundmodell des „Confirmation/ Disconfirmation-Paradigmas“ („C/D-Paradigma“), der gängigsten Konzeptualisierung des Konstruktes Kundenzufriedenheit: Die Bestätigung („confirmation“) der Erwartungen durch die tatsächlich wahrgenommene Leistung führt zur Zufriedenheit (Homburg & Stock-Homburg 2006). Unzufriedenheit resultiert demnach aus zu geringen Leistungen, zu geringer Wahrnehmung, zu hohen Ansprüchen oder einer Kombination aus allen dreien.

Die Kundenzufriedenheitsmessung ist bei der Definition der Qualitätsziele und der Bewertung der Zielerreichung insbesondere weicher, aber auch harter Angebotsmerkmale das wichtigste Instrument. Die DIN EN 13816 gibt hierzu Anhaltspunkte:

Abbildung 1: Qualitätskreis für Dienstleistungen im ÖPNV



Quelle: Eigene Darstellung nach DIN EN 13816:2002

Kundenzufriedenheit ist aus dieser Sicht keine statische Größe, sondern unterliegt in zweierlei Hinsicht einem Prozesscharakter: Zum einen können sich Wahrnehmungen und vor allem das Anspruchsniveau im Zeitverlauf ändern. Angebotsverbesserungen, die zum Zeitpunkt ihrer Einführung zu einer Erhöhung der Zufriedenheit geführt haben (z.B. Taktverdichtung), werden im Laufe der Zeit zum Standard und ziehen unter Umständen Wünsche nach weiteren Verbesserungen (Taktverdichtungen etc.) nach sich. Dies bedeutet aber auch, dass das Anspruchsniveau in gewissen Grenzen vom Anbieter beeinflussbar ist, indem unangemessen hohe Erwartungen relativiert werden. Zum anderen äußert sich der Prozesscharakter darin, dass Dienstleistungen in mehrere Phasen zerfallen, wobei sich die Zufriedenheit mit einer vorangegangenen Phase auf die Zufriedenheit mit der Qualität in späteren Phasen auswirken kann (Stauss & Seidel 2006).

An dieser Stelle sei jedoch bereits angemerkt, dass eine pauschale Bewertung, welche konkreten ÖPNV-Maßnahmen zu einer Erhöhung bzw. Verringerung der Kundenzufriedenheit führen, nicht ohne weiteres möglich ist. Diesbezügliche Ableitungen sind immer eng verknüpft mit den spezifischen Rahmenbedingungen vor Ort, wie z.B. dem jeweils vorliegenden „ÖPNV-Klima“.

Solch differenzierte Betrachtungen können allerdings nicht Gegenstand eines solchen Methodenleitfadens sein, da sie den inhaltlichen Rahmen bei weitem sprengen würden.

1.3 Zusammenhang von Kundenzufriedenheit und wirtschaftlichem Erfolg

Im Rahmen des diesem Leitfaden zugrunde liegenden Forschungsprojektes wurde eine *standardisierte Kundenzufriedenheitsbefragung* mit insgesamt 4.855 Interviews durchgeführt. Die Untersuchung war als Panel angelegt, im Abstand von einem Jahr wurden 1.702 Personen (von 3.153 in der ersten Erhebung) erneut zu den gleichen Themen befragt. Die Erhebung fand in drei Städten statt (Bonn, Tübingen und Ludwigshafen), in denen zwischen den beiden Erhebungswellen angebotsverändernde Maßnahmen im ÖPNV eingeführt wurden (siehe Abschnitt 3.2).

Eines der zentralen Ziele dieser Studie bestand darin zu untersuchen, ob es einen *Zusammenhang zwischen Kundenzufriedenheit und wirtschaftlichem Unternehmenserfolg* gibt. Auf Basis der Paneldaten wurde ein statistisches Modell zum Einfluss der Zufriedenheitsänderung auf die Veränderung der Kundenbindung - als Indikator des wirtschaftlichen Erfolgs - unter Kontrolle weiterer Merkmale entwickelt.

Als Hauptergebnis ergibt sich ein positiver Zusammenhang zwischen Zufriedenheit und Kundenbindung dergestalt, dass eine Veränderung (z.B. Erhöhung) der Globalzufriedenheit bei einer Person innerhalb eines Jahres um 1 Skalenpunkt (auf einer Skala von 1 bis 5) im Mittel zu einer Erhöhung des Kundenbindungsindex um rund 0,1 Skalenpunkte führt (bei Konstanzhaltung der übrigen Einflussmerkmale). Dieser Zusammenhang zwar nicht besonders stark ausgeprägt, aber statistisch signifikant. Dasselbe gilt für die Variable ÖPNV-Affinität: Auch die Veränderung der generellen Grundhaltung gegenüber dem ÖPNV („Die Benutzung des ÖPNV ist mir sympathisch“) wirkt sich somit auf die Kundenbindung aus.

Zusätzlich und unabhängig von den Paneldaten wurde der Zusammenhang zwischen Kundenzufriedenheit und Unternehmenserfolg auch auf der Aggregatebene, d.h. auf der Ebene der Verkehrsunternehmen bzw. -verbände analysiert. Hierzu wurden für Unternehmen, die zwischen 2003 und 2008 (mehrfach) am ÖPNV-Kundenbarometer (TNS-Infratest) teilgenommen haben, die entsprechenden Mittelwerte der Globalzufriedenheit mit bestimmten - vom VDV jährlich erhobenen - wirtschaftlichen Unternehmenskennziffern verknüpft. Unter Verwendung spezieller Regressionsmodelle, welche die Zeitreihenstruktur der Daten berücksichtigen, wurden Analysen für die Nachfragegrößen Ertrag (in 1.000 Euro), Anzahl Fahrgäste (in 1.000) und An-

zahl Personenkilometer (in 1.000) durchgeführt. Dabei zeigte sich im Hinblick auf den Unternehmensertrag und die Anzahl beförderter Fahrgäste (Unternehmensfahrten) ein signifikanter Zusammenhang mit der Kundenzufriedenheit.

Diese und weitere Ergebnisse sind im Detail und mit ausführlichen Erläuterungen in Kapitel 9 des Schlussberichts (siehe Bäumer, Hussinger und Pfeiffer 2010) dargestellt.

Die eben genannten Ergebnisse zeigen, dass dem Bereich Kundenzufriedenheit/Kommunikation eine wichtige Rolle für die weitere wirtschaftliche Entwicklung eines Unternehmens zukommt. Neben einer externen Beratung sichert vor allem eine feste personelle Verankerung dieses Bereiches im Unternehmen eine kontinuierliche Betreuung dieser Thematik und erleichtert die Entwicklung geeigneter Maßnahmen und - mindestens genauso wichtig - deren anschließende Umsetzung. Hierzu werden konkrete Praxistipps in Kapitel 5 dieses Leitfadens gegeben.

2 Überblick über die Methoden zur Kundenzufriedenheits- und Qualitätsmessung

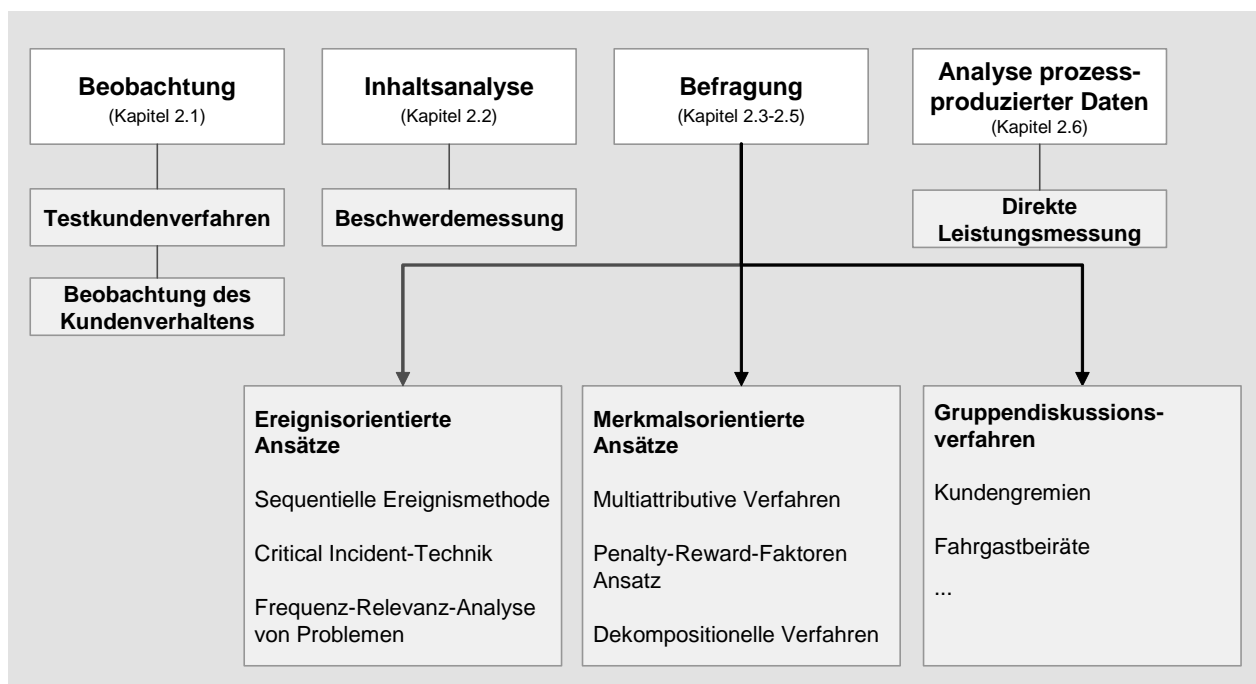
Im vorliegenden Kapitel werden zunächst die gängigen Verfahren zur Kundenzufriedenheits- und Qualitätsmessung beschrieben. Hierbei werden die betrachteten Verfahren auch mit kurzen Praxisbeispielen illustriert. In Kapitel 3 wird dann detailliert auf Methoden eingegangen, die auf Befragungen von Kunden basieren - insbesondere multiattributive Verfahren und die Critical Incident-Methode.

In der empirischen Sozialforschung werden als hauptsächliche Datenerhebungsverfahren

- Beobachtung
- Inhaltsanalyse
- Befragung und
- „nichtreaktive“ Verfahren, darunter die Analyse „prozessproduzierter“ Daten

unterschieden (vgl. z.B. Diekmann 2007). Wie im Folgenden zu sehen sein wird, bedient sich die Kundenzufriedenheitsforschung aller genannten Verfahren. Abbildung 2 zeigt zunächst einen systematisierenden Überblick über die zu behandelnden Methoden.

Abbildung 2: Überblick über die Methoden zur Kundenzufriedenheits-/ Qualitätsmessung



Es soll bereits an dieser Stelle darauf hingewiesen werden, dass die Wahl der Messmethode nicht völlig unabhängig vom Zweck der jeweiligen Untersuchung ist. Kundenzufriedenheitsmessungen zur Gestaltung von Besteller-Ersteller-Verhältnissen (Verkehrsverträge) unterliegen gegebenenfalls anderen Anforderungen als z.B. Anwendungen für rein betriebliche und marketinggestützte Optimierungen. Eine Bewertung der Einsatzbereiche und -grenzen der einzelnen Verfahren findet sich in Kap. 4.

2.1 Beobachtung

Die nicht-teilnehmende Beobachtung des Verhaltens von Kunden und Mitarbeitern durch geschulte (externe) Forscher dient hauptsächlich dazu, Mängel im Prozess der Dienstleistungserstellung zu erkennen (z.B. Warteschlangen an Fahrschein-Verkaufsstellen etc.). Da es aber oftmals nicht möglich ist, diese Beobachtung vollkommen verdeckt durchzuführen, kann es dabei zu Beobachtungseffekten, also ungewollten Verhaltensänderungen in der untersuchten Situation, kommen. Darüber hinaus ist der Prozess der Dienstleistungserstellung häufig nur in Teilen beobachtbar. Wenngleich die Objektivität der gewonnenen Daten aufgrund der Neutralität und der Ausbildung der Beobachter oft gewährleistet werden kann, besteht das größte Problem darin, dass der Schluss von beobachtetem Verhalten auf die subjektiv wahrgenommene Zufriedenheit der Kunden mit großen Unsicherheiten behaftet ist.

Auch der Einsatz von Schein- oder Testkunden, das so genannte „Silent-Shopping“ (auch „Mystery Shopping“), kann als Beobachtungsverfahren (teilnehmend, meist verdeckt) klassifiziert werden. Dabei lassen sich drei Arten von Testkunden unterscheiden: Zum ersten Mitarbeiter des Anbieters, die nicht verdeckte Qualitätsprüfungen nach standardisierten Kriterien vornehmen, wobei hier auch Bereiche geprüft werden (können), die außerhalb der Sichtbarkeit durch den Kunden liegen. Zum zweiten externe Experten, die nicht dem Unternehmen des Anbieters angehören, sich als normale Kunden geben und die umfangreiche Erfahrungen bei solchen Prüfungen haben. Drittens werden Silent Shopping-Verfahren auch mit echten Kunden durchgeführt, die im Auftrag des Anbieters verdeckte Testkäufe durchführen (vgl. Schnippe 2000; VDV 2002).

Generell hängt die Qualität der Ergebnisse solcher Verfahren stark von den Fähigkeiten der Testkunden ab. So wird der Anbieter in der Regel an der Bewertung der Qualität „normaler“ Dienstleistungsinteraktionen interessiert sein (und nicht an außergewöhnlichen oder kritischen Ereignissen). Insofern müssen die Tester in der Lage sein, solche normalen Situationen zu simulieren und zu bewerten. Darüber hinaus müssen sie die Fähigkeit haben, die gewonnenen Eindrücke im Nachhinein unverzerrt zu dokumentieren.

Wie bei der nicht-teilnehmenden Beobachtung stellt sich auch bei Testkundenverfahren die Frage, ob von der Wahrnehmung professioneller Beobachter auf die Kundenwahrnehmung und damit auf die -zufriedenheit geschlossen werden kann, zumal oft nicht alle für den Kunden wichtige Merkmale beim Testkauf berücksichtigt werden können (vgl. Stauss 2000; Meffert & Bruhn 2008). Da beim Einsatz „echter“ Kunden als Tester die Beurteilung der gemachten Erfahrungen tatsächlich aus Kundensicht erfolgt, ist diese Vorgehensweise am ehesten auf das Konzept der Kundenzufriedenheit zugeschnitten. Allerdings bestehen hier durchaus auch Verzerrungsgefahren, insbesondere wenn die Testkunden unzureichend geschult sind oder nicht über einschlägige Branchenkenntnisse verfügen.

Praxisbeispiele:

Die üstra Hannoversche Verkehrsbetriebe AG führt seit Januar 2006 in den eigenen Bussen und Stadtbahnen Qualitätstests mittels Mystery Shopping durch. Zu diesem Zweck wurden zwölf Studenten für den Einsatz als Testkunden geschult. Die Tester sind monatlich jeweils auf sechs so genannten Routen unterwegs und beobachten Service und Sauberkeit in den Verkehrsmitteln. Sie stellen dem Personal typische Fragen und beurteilen Fahrweise, Kundenfreundlichkeit/Fahrgastorientierung, Fachkompetenz, Pünktlichkeit und äußeres Erscheinungsbild des Fahrers. Diese Kriterien werden in insgesamt 22 Beurteilungspunkte unterteilt und von den Testern nacheinander abgearbeitet. Weiter wird ein subjektiver Gesamteindruck vom Tester notiert, welcher jedoch nicht in die Gesamtwertung einfließt. Die Checklisten werden anschließend vom Qualitätsmanagement intern ausgewertet und als Punktesystem gespeichert. Die Daten werden anschließend in unterschiedlichen Unternehmensbereichen der üstra genutzt (vgl. Lindenberg, Prella, Keese & Uhlenhut 2007).

Die Berliner Verkehrsbetriebe (BVG) setzen seit 1996 regelmäßig Kunden als Qualitätstester ein. Die Kunden füllen monatlich oder mehrmals jährlich unentgeltlich Fragebögen zum Thema Leistungs- und Serviceangebot der BVG aus. Teilweise werden die Tester zusätzlich eingeladen, um eine Bewertung direkt am Objekt vorzunehmen. Zum Erhebungsprogramm gehören Qualitätsmerkmale, Sauberkeitsaspekte, Mitarbeiterbewertung, Tarifanpassung, linienbezogene Bewertungen, Informationskette, Kundenzeitschrift, Verkauf von Zeitungen in den Fahrzeugen und die Akzeptanz des Callcenters. Auf Grund der Handlungsempfehlungen wurden entsprechende qualitäts- und attraktivitätssteigernde Maßnahmen, wie die Anzahl und Ausstattung der Haltestellen mit Wartehallen sowie Anzahl und Lage von fahrplanmäßigen Anschlüssen, eingeführt. Die Gruppe wurde nach einer Anzeige in der Kundenzeitschrift aufgebaut und bestand zunächst aus 403 Testkunden. Die Gruppe der Qualitätstester weist im Vergleich zu einer repräsentativen Stichprobe der Berliner Bevölkerung Unterschiede auf. So besteht die Gruppe beispielsweise beinahe nur aus Stammkunden, was auch nur ein Vergleich mit diesen Kunden aus der repräsentativen Bevölkerungsstichprobe zulässt. Die BVG erhebt neben den soziode-

mographischen und individuellen Merkmalen die Globalzufriedenheit mit der BVG und allen erhobenen Qualitätsmerkmalen. Daraus ergaben sich fünf Gruppen von Testkunden. Die erste Gruppe, 10 % der Qualitätstester im Alter bis 20 Jahre, bewerteten die BVG besonders positiv. Die zweite Gruppe, 23 % der Qualitätstester im Alter zwischen 20 und 39 Jahren, die einen Pkw zur Verfügung haben, sind kritische Kunden. Die dritte Gruppe (15 % der Qualitätstester), vorwiegend über 60 Jahre alt, sind „Nörgler“. Die anderen beiden Gruppen liegen zwischen diesen Gruppen und im Altersbereich zwischen 30 und 59 Jahren. Die Qualitätstester arbeiten ehrenamtlich, um eine Abhängigkeit von der BVG und somit eine zu positive Bewertung zu vermeiden. Dennoch müssen die Testkunden durch materielle und ideelle Incentives umworben werden. So ist den Qualitätstestern beispielsweise besonders wichtig, dass ihnen bei Veranstaltungen Vorstandsmitglieder persönlich zur Kontaktaufnahme zur Verfügung stehen. Da die Tester sehr engagiert sind, entsteht ein intensiver Dialog mit der BVG, was einen erheblichen Betreuungsaufwand mit sich bringt. Diesen Aufwand hält man allerdings für gerechtfertigt, da das Qualitätskonzept ein wesentliches Element der Kundenorientierung der BVG ist. Eine Wirkung dieser Maßnahme ist auch die Verbesserung des Unternehmensimages (vgl. Buschmann 2000).

Weitere Anwendungsbeispiele:

- Region Hannover: Mystery Shopping im Rahmen des vertraglichen QS (Felscher-Suhr, Mattern & Schreckenber 2005)
- Ergebnisse eines Mystery Shoppings bei der Straßenbahn Herne-Castrop-Rauxel (Lüttenberg, Person & Felscher-Suhr 2004)

2.2 Inhaltsanalyse

Die systematische Erfassung und Auswertung schriftlich oder mündlich vorgetragener Rückmeldungen von Kunden - negativer wie positiver Art - kann als inhaltsanalytisches Verfahren aufgefasst werden. Vorteile der **Beschwerdeanalyse** können darin gesehen werden, dass sich Beschwerden immer auf einen konkreten Sachverhalt beziehen, stets aktuell sind und darüber hinaus eine für das Unternehmen relativ kostengünstige Informationsquelle zur Identifikation von betrieblichen Schwächen und Marktchancen darstellen¹. Der Nachteil besteht darin, dass Kundenbeschwerden hinsichtlich der Beschwerdeführer und der dabei vorgetragenen Probleme

¹ In der Analyse und vor allem der adäquaten Bearbeitung von Beschwerden durch das Unternehmen kann auch ein Instrument zur Kundenbindung gesehen werden. Untersuchungen zeigen, dass Kunden, deren Beschwerde für sie zufriedenstellend bearbeitet wurde, eine höhere Zufriedenheit aufweisen als solche, die sich nicht beschwerten.

nicht repräsentativ sind, da sich nur ein geringer Anteil von unzufriedenen Kunden zu einer Beschwerde entschließt. So führen nur etwa 2-4 % aller Problemereignisse zu einer schriftlichen Beschwerde (Albrecht & Bradford 1990). Ein weiteres Problem ergibt sich oft auch dadurch, dass mündliche Rückmeldungen an „normale“ Mitarbeiter des Unternehmens (also nicht an speziell zu diesem Zweck eingerichtete Stellen) nicht weitergegeben werden. Dies liegt zum Teil an fehlendem Wissen, wem die Beschwerde zugeleitet werden muss, an nicht vorhandenen Anreizen zur Weitergabe und natürlich auch am Bestreben der Mitarbeiter, sich nicht selbst in ein schlechtes Licht zu rücken (vgl. Stauss 2000; Hentschel 1992).

Praxisbeispiele:

In einem Pilotprojekt entwickelte der Stadtverkehr Hürth in Zusammenarbeit mit IT-Spezialisten ein spezielles System für einen internetgestützten Kundendialog. Das System erleichtert sowohl die Eingabe als auch die Verarbeitung von Beschwerden, Anregungen/Anfragen und Mitteilungen. Dabei entfallen viele Vorgänge in der Bearbeitung der Kundenbeschwerden und Anregungen, so dass die Mitarbeiter schneller und effizienter Kundeneingaben bearbeiten können (Wiethüchter, Eidam, Knoch & Dallmeyer 2007).

Um die Zufriedenheit der Kunden mit dem öffentlichen Nahverkehr zu erhöhen, führte der Nordhessische Verkehrsverbund 2006 eine Fünf-Minuten-Garantie ein. Die Fünf-Minuten-Garantie beinhaltet fünf Versprechen in den Bereichen Pünktlichkeit, Anschlusssicherheit, saubere Fahrzeuge und Haltestellen, guter Service sowie präzise Information und Freundlichkeit des Personals. Um diesbezügliche Beschwerden für den Kunden möglichst einfach zu gestalten, stehen alle üblichen Kommunikationswege wie das Internet, schriftliche und persönliche Beschwerde in allen NVV-Kundenzentren und NVV-InfoPoints zur Verfügung. In diesem Zusammenhang wurde ein internet-basiertes Beschwerdemanagementsystem entwickelt, mit dem die Bearbeitung der Beschwerden wesentlich einfacher geworden ist und dem Unternehmen konkrete Ansatzpunkte für Verbesserungen liefert. Mit Hilfe des neuen Systems können entsprechende Schwachstellen sofort erkannt und automatisch an die zuständige Stelle weitergeleitet werden. Durch das Beschwerdemanagementsystem entsteht eine hohe Transparenz, die eine schnelle Mängelanalyse und im Idealfall eine schnelle Mängelbeseitigung ermöglicht (vgl. Kepper 2007).

2.3 Kundenzufriedenheitsmessung mittels Befragung: Ereignisorientierte Verfahren

Bei der Messung der Kundenzufriedenheit über die Einzelbefragung von Kunden kann man grundsätzlich zwischen ereignisorientierten und merkmalsorientierten Verfahren unterscheiden.

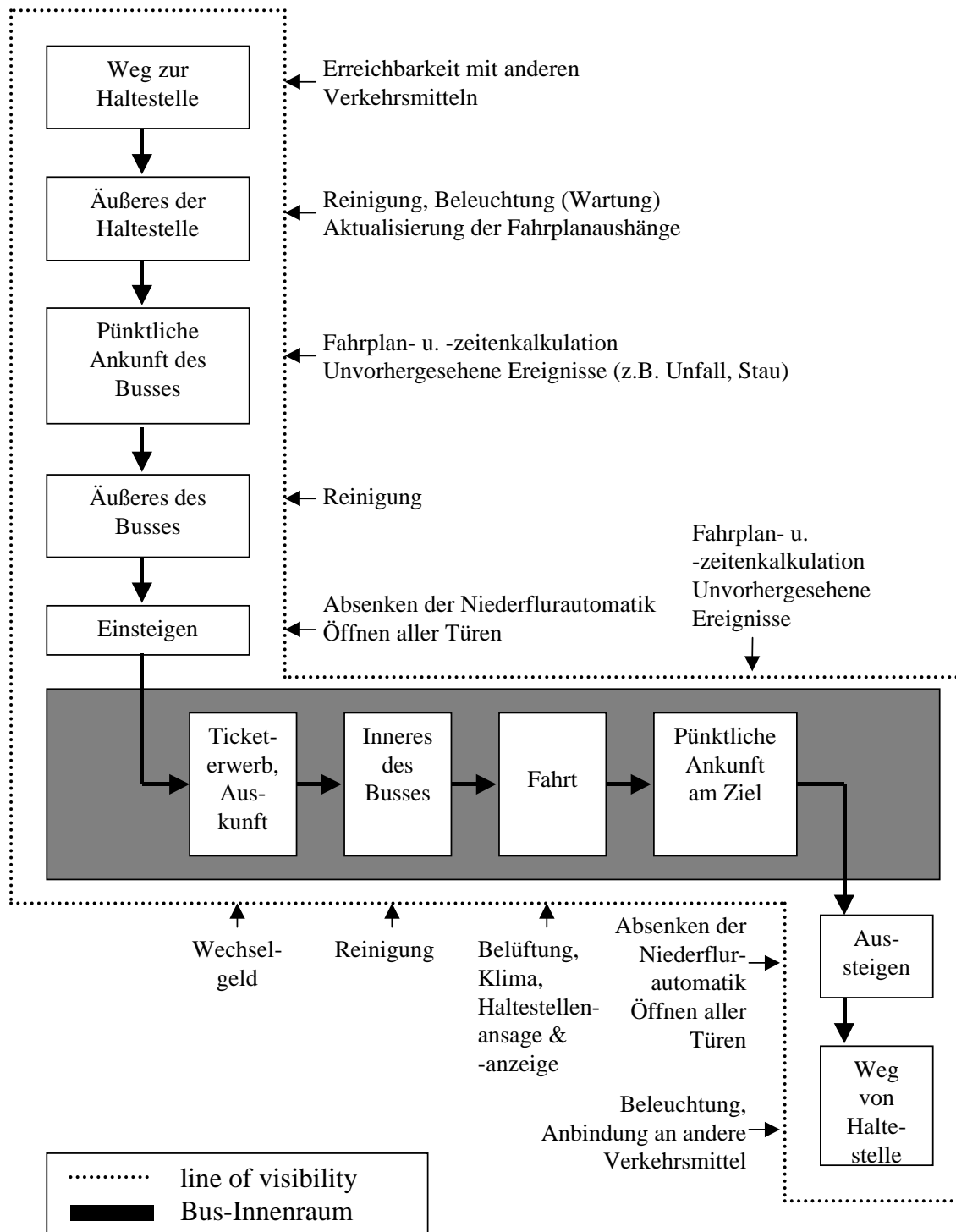
Die unter dem Etikett „ereignisorientierte Verfahren“ zusammengefassten Ansätze sind insbesondere für die Messung der Kundenzufriedenheit bei Dienstleistungen, wie z.B. dem ÖPNV, relevant. Ein konstituierendes Element von Dienstleistungen ist ja die Beteiligung des Kunden am Prozess der Leistungserstellung. Demgemäß beziehen sich diese Verfahren speziell auf die Momente (Ereignisse), in denen der Kunde mit dem Dienstleister bzw. Elementen des Dienstleistungsangebots in Kontakt kommt (sog. „Augenblicke der Wahrheit“). Dabei wird davon ausgegangen, dass Zufriedenheit bildende Erfahrungen als Ereignisse mit einem konkreten räumlichen und zeitlichen Bezug wahrgenommen werden (Stauss 2000; Stauss & Seidel 2006). Ereignisorientierte Verfahren dienen somit vor allem dem Zweck, Informationen darüber zu erhalten, welche Phasen des Dienstleistungsprozesses einer besonderen Aufmerksamkeit des Qualitätsmanagements bedürfen (Meffert & Bruhn 2008).

2.3.1 Sequentielle Ereignismethode

Das erste hier darzustellende Verfahren, das dem Rechnung trägt, ist die Methode der sequentiellen Ereignisse, die insbesondere zur Analyse „gewöhnlicher“ Kundenkontakte geeignet ist. Dabei wird so vorgegangen, dass zunächst der Dienstleistungsprozess in seine einzelnen Phasen („Kundenkontaktpunkte“) zerlegt wird. Hier ist darauf zu achten, dass diese Gliederung aus der Perspektive des Kunden erfolgt. Das Ergebnis dieses ersten Analyseschrittes wird üblicherweise in Form eines grafischen Ablaufschemas - auch „Blueprint“ genannt - dargestellt. Die Ausarbeitung eines solchen Blueprints kann durch die Ermittlung der Auftretenshäufigkeit der identifizierten Kontaktpunkte ergänzt werden. Dies kann z.B. über die Auswertung betriebsinterner Informationen oder auch über gesonderte Erhebungen geschehen (vgl. Stauss 2000, S. 329).

Abbildung 3 zeigt ein solches Ablaufschema am Beispiel einer Fahrt mit einem Linienbus. Danach beginnt der Dienstleistungsprozess mit dem Weg zur Einstiegshaltestelle und endet mit dem Verlassen der Ausstiegshaltestelle. Dabei sollten allerdings auch mögliche Verzweigungen im Ablauf der Dienstleistung berücksichtigt werden, so kann im Beispiel der Fahrscheinerwerb beim Fahrer oder am Automaten stattfinden bzw. - wegen Zeitkartenbesitz - auch ganz entfallen.

Abbildung 3: Beispiel eines Blueprinting der Dienstleistung „Personenbeförderung mit Linienbus“



Quelle: Schnippe 2000, S. 181

Im Schaubild sind auch die Komponenten des Dienstleistungsprozesses abgebildet, die den eigentlichen Kontaktpunkten vorgelagert sind (z.B. Reinigung des Transportgefäßes). Diese werden vom Kunden zwar in aller Regel nicht unmittelbar erfahren, üben aber dennoch einen

Einfluss auf die Qualitätswahrnehmung aus. Die Trennung dieser beiden Komponenten erfolgt über die so genannte „line of visibility“.

Anhand eines solchen Ablaufschemas wird der Befragte in einem persönlichen Interview durch den erlebten Dienstleistungsprozess geführt, und es wird mittels offener Fragen nach dem wahrgenommenen Ablauf, den Empfindungen sowie positiven bzw. negativen Beurteilungen an den einzelnen Kundenkontaktpunkten gefragt (Stauss 2000). Man erhält damit einen sehr detaillierten Einblick in das Dienstleistungserleben des Kunden, der allerdings aufgrund der Orientierung an einem Ablaufschema (sog. „aided recall-Aufgabe“) immer auch einen gewissen Teil an irrelevanten Informationen enthält (vgl. Hentschel 1992).

Aus diesem Grund bietet es sich an, die geschilderte qualitative Vorgehensweise durch eine quantitative zu ergänzen. Hier wird zusätzlich am Ende des Interviews die Zufriedenheit mit dem gesamten Dienstleistungsprozess („Transaktionszufriedenheit“), und nach jedem Kontaktpunkt die zugehörige „Episodenzufriedenheit“ - jeweils in Form geschlossener Fragen - erhoben. Hieraus können Informationen über den Einfluss der einzelnen Episodenzufriedenheiten auf die Gesamtbewertung und - im Sinne einer prozessualen Zufriedenheitsermittlung (vgl. hierzu Stauss & Seidel 2006) - über die Zufriedenheitsdynamik, also deren Veränderung im Verlauf der Dienstleistung, gewonnen werden.

Praxisbeispiel:

Die Verkehrs- und Aktiengesellschaft Nürnberg führt seit vielen Jahren Kundenzufriedenheitsuntersuchungen durch und greift hierbei auf einen Methodenmix zurück, der u.a. auch die sequentielle Ereignismethode abdeckt. Hierbei wird die Zufriedenheit mit den einzelnen Gliedern eines konkreten ÖPNV-Weges offen erhoben. Durch die freie Abfrage der Zufriedenheit sollen Verzerrungen durch die Vorgabe verschiedener Merkmale vermieden werden (vgl. Dahlmann-Resing, Gebauer & Klodner 2004).

2.3.2 Critical Incident-Technik

Im Gegensatz zur sequentiellen Ereignismethode hebt die Critical Incident-Technik nur auf besondere Vorfälle, die als außergewöhnlich positiv oder negativ empfunden werden, ab (z.B. defekter Fahrscheinautomat, lange Wartezeit an der Verkaufsstelle etc.). Mittels standardisierter, offener Fragen wird der Kunde veranlasst, an ein besonders gutes oder schlechtes Erlebnis mit dem Dienstleister zurückzudenken und dies möglichst detailliert zu beschreiben (Wann war das? Welche Umstände führten dazu? Verhalten und Aussagen der Mitarbeiter? Was geschah genau, dass der Kontakt als (nicht) zufriedenstellend empfunden wurde (vgl. Meffert & Bruhn 2008).

Die hierbei erzielbaren Informationen werden in der Regel als sehr relevant eingestuft, was die Identifikation von Stärken und vor allem Schwächen beim Ablauf der Dienstleistung anbelangt, da aufgrund der ungestützten Erinnerung („unaided recall“) die Kunden vorwiegend solche Ereignisse berichten, die ihnen im Gedächtnis haften geblieben und somit von besonderer Bedeutung für sie sind. Auf die Critical Incident-Technik wird in Kap. 3.4 detailliert eingegangen.

Praxisbeispiele:

Im Juli und August 1998 wurden an zentralen Nahverkehrsknotenpunkten in Berlin, Greifswald, Hamburg, Lübeck, Magdeburg, Neubrandenburg, Rostock und Stralsund 400 Personen befragt. Bei den Fragen wurde die Critical Incident-Technik angewendet. Die Personen wurden gebeten, insgesamt drei Fragen zu beantworten. Die erste gestellte Frage bezog sich auf ein konkretes Ereignis: „Hatten Sie schon einmal ein besonders gutes (schlechtes) Erlebnis mit dem öffentlichen Personennahverkehr von *Name der Stadt*? Bitte schildern Sie es mir.“ Weitere gestellte Fragen waren: „Was macht Ihrer Meinung nach einen besonders guten (schlechten) öffentlichen Personennahverkehr aus?“ „Was ist Ihnen am öffentlichen Personennahverkehr besonders wichtig?“ Um Reihenfolgeeffekte zu vermeiden, wurde die Polung (gut/schlecht) der ersten beiden Fragen systematisch variiert. Im Folgenden wurde versucht, die Kundenaussagen zu kategorisieren. Es wurde dabei deutlich, dass die Pünktlichkeit gefolgt von Sicherheit und einer hohen Taktfrequenz für die Kunden am wichtigsten sind (vgl. Schnippe 1999).

Die Pinneberger Verkehrsgesellschaft mbH (PVG) führte Ende 1999 eine Fahrgastbefragung durch, die neben einer Einteilung der Kunden in Gruppen auch eine differenzierte Analyse der Leistungen liefern sollte. Mittels der Methode der kritischen Ereignisse wurden Aussagen von 248 Befragten, die von insgesamt 396 persönlich erlebten Ereignissen berichteten, analysiert. Die Aussagen wurden kodiert und verschiedenen Kategorien zugeordnet. 43 Prozent aller Erlebnisse und Aussagen betreffen das Verhalten des Personals. 42 Prozent der negativen Aussagen betreffen das Leistungsangebot, wobei sich hier wiederum 40 Prozent auf mangelnde Pünktlichkeit beziehen. Von den insgesamt 20 gebildeten Faktoren wurden zehn mit einem signifikanten Einfluss auf die Gesamtzufriedenheit identifiziert (vgl. Probst, Kahrs & Rieschick-Dziabas 2001).

Sowohl die sequentielle Ereignismethode als auch die Critical Incident-Technik erfordern aufgrund ihrer qualitativen Ausrichtung einen hohen Erhebungs- und Auswertungsaufwand (Inhaltsanalyse von Gesprächsprotokollen bzw. offenen Fragen), was oftmals zur Folge hat, dass nur sehr kleine Stichproben realisiert werden können (vgl. Stauss 2000). Dies gilt auch für Methoden, in denen die Befragungsperson lediglich ein Grobthema gestellt bekommt und zu diesem dann völlig frei, also ohne Rückgriff auf einen Fragebogen oder Leitfaden, ihre Erlebnisse berichten soll („story telling“ oder narratives Interview).

2.3.3 Frequenz-Relevanz-Analyse von Problemen (FRAP)

Bei der FRAP handelt es sich um ein quantitatives Verfahren der ereignisorientierten Messung. Vor allem für die Ableitung von Maßnahmen ist es wichtig, Informationen über die Häufigkeit des Auftretens (Frequenz) und die vom Kunden empfundene Bedeutung (Relevanz) von negativen Erlebnissen mit dem Dienstleister zu gewinnen. Grundlage hierfür ist eine Problemliste, die z.B. über die beiden zuletzt besprochenen Verfahren generiert werden kann. Bei jedem Problem soll dann der Kunde mittels eines standardisierten Fragebogens angeben, ob dieses bei ihm aufgetreten ist (Frequenz) und wenn ja, wie groß seine Verärgerung darüber ist und wie er auf dieses Problem reagiert hat bzw. plant, darauf zu reagieren (z.B. mit Anbieterwechsel oder dem Weitererzählen des Problems). Diesen Reaktionsformen werden dann Punktwerte zugeordnet und mit den Skalenwerten zum Ausmaß der Verärgerung multiplikativ zu einem Relevanzwert verknüpft. Somit liegen für jedes vorgelegte Problem Informationen zu dessen Häufigkeit und der empfundenen Bedeutsamkeit vor, die dann in Form von zweidimensionalen Diagrammen oder über die Bildung von Problemindizes analysiert werden können (Stauss 2000). Es ist klar, dass die Tauglichkeit dieses Verfahrens in entscheidendem Maße von der Vollständigkeit der zur Beantwortung vorgelegten Problemliste abhängt. Als kritisch ist darüber hinaus die etwas willkürliche Punktbewertung der Kundenreaktionen einzuschätzen.

Praxisbeispiele:

Hinweise auf konkrete Anwendungsbeispiele dieser Methode im ÖPNV-Bereich konnten bisher nicht gefunden werden. Die FRAP wird derzeit hauptsächlich zur Analyse der Servicequalität von Dienstleistungen oder der Qualität von internen Geschäftsprozessen im Bereich der Banken, Bibliotheken oder Kfz-Dienstleistungen eingesetzt.

Anmerkung

Gelegentlich werden die eben vorgestellten ereignisorientierten Verfahren auch unter dem Namen Kontaktpunktanalyse zusammengefasst. Dabei werden in einem ersten Schritt möglichst alle Kundenkontaktsituationen erfasst (z.B. in Form eines Blueprints; vgl. oben) und ggf. quantifiziert. Danach erfolgt die qualitative Kontaktpunktanalyse in Bezug auf gewöhnliche (sequentielle Ereignismethode oder auch Beobachtung) oder kritische (Critical Incident-Technik oder auch Beschwerdeanalyse) Ereignisse. Der dritte Schritt schließlich umfasst die quantitative Analyse insbesondere der kritischen Ereignisse mittels einer Frequenz-Relevanz-Analyse von Problemen.

2.4 Kundenzufriedenheitsmessung mittels Befragung: Merkmalsorientierte Verfahren

Während bei den eben besprochenen ereignisorientierten Verfahren kontextgebundene Episoden und Geschichten die Grundeinheit der Analyse bilden, findet bei merkmalsorientierten Verfahren die Qualitätsbewertung kontextungebunden auf der Basis allgemeiner Begriffe (Merkmale, z.B. Pünktlichkeit der Busse, Sauberkeit in den Fahrzeugen, Höflichkeit der Busfahrer usw.) statt.

2.4.1 Multiattributive Ansätze

Bei der Gruppe der multiattributiven Verfahren handelt es sich um die am weitesten verbreitete Methode der Kundenzufriedenheitsmessung, die z.B. auch im Rahmen des Deutschen Kundenmonitors (vormals Deutsches Kundenbarometer) angewandt wird (vgl. Meyer & Dornach 1995). Die Logik der multiattributiven Messung liegt auch dem sog. „SERVQUAL-Ansatz“ zugrunde, mit dem das Ziel verfolgt wurde, ein branchenübergreifendes Instrument zur Messung der Dienstleistungsqualität zu entwickeln (Parasuraman, Zeithaml & Berry 1988).

Der Kerngedanke der multiattributiven Verfahren besteht darin, dass sich ein Gesamtqualitätsurteil aufgrund einer spezifischen Kombination von Einzeleindrücken über Qualitätsmerkmale bildet. Die Wurzeln der multiattributiven Verfahren liegen in der sozialpsychologischen Einstellungsforschung, und zwar bei den so genannten Wert-Erwartungs-Modellen. Danach bildet sich die Einstellung gegenüber einem Objekt (z.B. ÖPNV) über die subjektive Wahrscheinlichkeit bezüglich des Vorliegens bestimmter Eigenschaften des Einstellungsobjektes (z.B. sauber, sicher, pünktlich usw.) und der Bewertung dieser Eigenschaften (Wichtigkeit von Sauberkeit, Sicherheit, Pünktlichkeit etc.). Formal kann dies folgendermaßen dargestellt werden (vgl. Rosenberg 1956):

$$E_{ij} = \sum_k x_{ijk} \cdot y_{ik}$$

Dabei ist E_{ij} die Einstellung der Person i gegenüber Objekt j ,

x_{ijk} die subjektive Wahrscheinlichkeit von Person i , dass Objekt j die Eigenschaft k besitzt und

y_{ik} die Wichtigkeit der k -ten Eigenschaft für Person i .

Zentral ist dabei die multiplikative Verknüpfung der Wahrscheinlichkeits- und Bewertungskomponente pro Eigenschaft² und die anschließende Summation über alle Eigenschaften³. Die Einstellung setzt sich somit aus einer Vielzahl von einzeln bewerteten Eigenschaften zusammen, weshalb diese Ansätze auch als kompositionelle Verfahren bezeichnet werden. Nicht nur zur Kontrolle der Validität ist es sinnvoll, in der Befragung ein separates Globalurteil zu erheben. Auf die multiattributiven Verfahren wird in Kap. 3.3 detailliert eingegangen.

Praxisbeispiele:

Eines der bekanntesten Anwendungsbeispiele für eine multiattributive Messung der Zufriedenheit ist das ÖPNV-Kundenbarometer. Das ÖPNV-Kundenbarometer wird seit 1999 jährlich erstellt, Auftraggeber sind verschiedene Verkehrsunternehmen und -verbände. Dabei werden standardisierte Untersuchungen mit kleinen Stichproben in eng umgrenzten Regionen durchgeführt. Innerhalb des Befragungsgebietes wird eine Zufallsauswahl vorgenommen und mit den entsprechenden Zielpersonen ein telefonisches Interview geführt. Als Mindeststichprobengröße wurden 250 Personen festgelegt, um repräsentative Aussagen zu gewährleisten. Zur Messung der Kundenzufriedenheit wird eine asymmetrische 5er-Skala verwendet. Die Grundlage des Auswertungsmodells bildet eine ausführliche Liste von Leistungsmerkmalen. In einem ersten Schritt erfolgt eine Beschreibung der Kundenzufriedenheit mit den entsprechenden Leistungsmerkmalen. Im zweiten Schritt werden die Werte der beteiligten Unternehmen hinsichtlich der Kundenzufriedenheit miteinander verglichen. Sowohl die Ergebnisse der Globalzufriedenheit als auch die Ergebnisse für die einzelnen Leistungsmerkmale werden den Unternehmen zur Verfügung gestellt, damit jedes Unternehmen die Möglichkeit hat, sich mit den Beteiligten zu vergleichen, die ähnliche Strukturen haben. Das Kundenbarometer ist ein Instrument, mit dem auch kleinere Unternehmen die Möglichkeit erhalten, sich einen Überblick über die Kundenzufriedenheit ihrer Fahrgäste zu verschaffen (vgl. Isfort 2000).

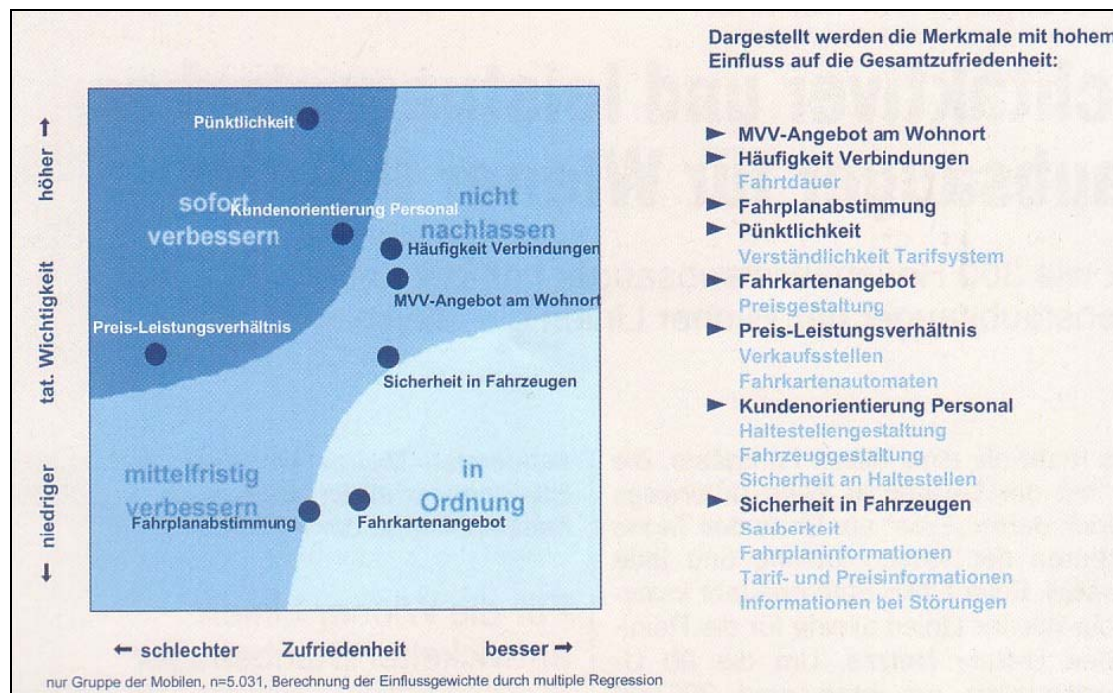
Der MVV führt seit 1995 ca. alle fünf Jahre Studien zur Gewinnung von Daten zur MVV-Marktsituation durch. Ziel der letzten Studie war primär die Erhebung von Daten zur Kundenstruktur und zum Kundenverhalten, um effektivere Angebotsverbesserungen zu erreichen und ein zielgruppenorientiertes Marketing einzuführen. In diesem Zusammenhang wurde auch die Zufriedenheit mit den Leistungsmerkmalen des ÖPNV abgefragt und im Rahmen einer so genannten Portfolio-Analyse aus der Relation zwischen Zufriedenheit und Wichtigkeit entspre-

² Hierbei muss die Unabhängigkeit der beiden Komponenten voneinander unterstellt werden.

³ Hierbei muss die isolierte Wirksamkeit eines jeden Merkmals für die Globaleinstellung vorausgesetzt werden.

chender Handlungsbedarf abgeleitet (vgl. Strasser, Haller & Krietemeyer 2008). Wie eine Portfolio-Analyse im ÖPNV-Bereich aussehen kann, ist in der nachfolgenden Abbildung dargestellt.

Abbildung 4: Portfolio zur Analyse der ÖPNV-Kundenzufriedenheit



Quelle: Strasser, Haller & Krietemeyer 2008

2.4.2 Penalty-Reward-Faktoren-Ansatz

Während die Logik der eben besprochenen multiattributiven Verfahren impliziert, dass eine schlechte Bewertung eines Merkmals durch eine positive Beurteilung eines anderen Merkmals in Bezug auf das Gesamturteil kompensiert werden kann (deshalb „kompensatorische Modelle“), wird bei der Penalty-Reward-Faktoren-Analyse unterstellt, dass „die Qualitätsmerkmale bestimmte minimal tolerierbare Niveaus aufweisen müssen“ (Hentschel 1992, S. 114). Es wird davon ausgegangen, dass es einerseits Merkmale (Faktoren) gibt, die - falls die diesbezüglichen Zufriedenheiten abnehmen - zu einer überdurchschnittlichen Abnahme und - falls die Zufriedenheiten zunehmen - zu keiner deutlichen Steigerung der Gesamtzufriedenheit führen (Penalty-Faktoren). Dem gegenüber stehen Merkmale (Reward-Faktoren), bei denen eine Steigerung der Zufriedenheit ein Anwachsen der Globaleinschätzung, ein Absinken der Zufriedenheit aber nicht eine Reduktion des Gesamturteils nach sich zieht (vgl. Meffert & Bruhn 2008). Die Auswertung erfolgt über eine multiple Regression mit dem Gesamtqualitätsurteil als abhängige Größe und den in dummy-(0/1-)Variable zerlegten merkmalsbezogenen Zufriedenheiten als Prädiktoren.

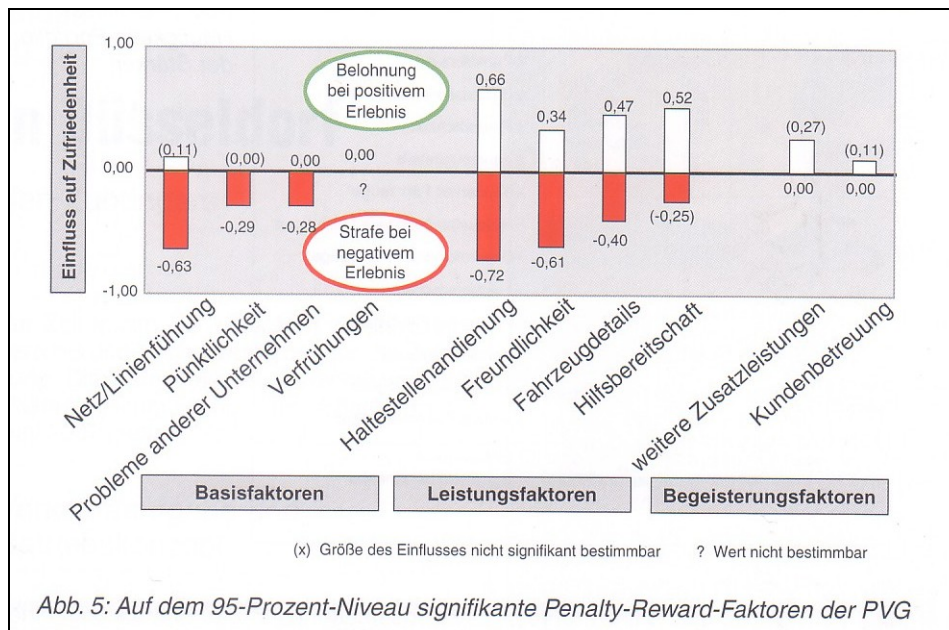
Das Ziel des Verfahrens besteht darin, speziell die Penalty-Faktoren - also die Mindestanforderungen, die im Fall der Nicht-Erfüllung bei den Kunden Unzufriedenheit hervorrufen - zu identifizieren, da sich hieraus gezielte Strategien ableiten lassen. In empirischen Untersuchungen gelingt dies jedoch nicht immer. Oft ist es vielmehr so, dass Merkmale in beide Richtungen, also sowohl als Penalty- wie auch als Reward-Faktor, wirken (vgl. z.B. Bogner & Plotegher 1997).

Praxisbeispiele:

In der oben erwähnten Erhebung der Pinneberger Verkehrsgesellschaft mbH (PVG) Ende 1999 wurden die mittels Critical Incident-Technik erzeugten Variablen einer Penalty-Reward-Faktoren-Analyse unterzogen. Hier werden konkrete Ereignisse in das Qualitätsmodell von Kano, unterteilt nach Basisanforderungen, Leistungsanforderungen und Begeisterungsanforderungen sortiert.

Nach der Einteilung wird deutlich, dass sich Basisfaktoren, wie Netz/ Linienführung und Pünktlichkeit zwar im negativen Fall signifikant auf die Zufriedenheit auswirken, im positiven Fall jedoch nicht honoriert werden. Auch negative Erfahrungen mit anderen Verkehrsunternehmen, die nichts mit der PVG zu tun haben, strahlen negativ auf die Kundenzufriedenheit aus. Es wird ebenfalls deutlich, dass die vier Faktoren Verfrühung, Haltestellenandienung, Freundlichkeit und Hilfsbereitschaft einen großen Einfluss auf die Zufriedenheit der Kunden haben. Diese Leistungsfaktoren haben im negativen und im positiven Sinn einen bedeutenden Einfluss auf die Zufriedenheit. Weitere Zusatzleistungen, wie etwa der Zeitungsverkauf im Fahrzeug oder das Schmücken der Fahrzeuge zu besonderen Anlässen und Kundenbetreuung zählen zu den Begeisterungsfaktoren, mit positivem Einfluss auf die Kundenzufriedenheit (vgl. Probst, Kahrs & Rieschick-Dziabas 2001).

Abbildung 5: Beispiel für die Anwendung des Penalty-Reward-Faktoren-Ansatzes



Quelle: Probst, Kahrs & Rieschick-Dziabas 2001

In einer Untersuchung von Bogner & Plotegher (1997) wurde im Rahmen einer bundesweiten telefonischen Befragung von ÖPNV-Kunden im Zeitraum 1996/1997 u.a. die Wichtigkeit und die Wirkungsrichtung einzelner Angebotsmerkmale des ÖPNV im Rahmen einer Penalty-Reward-Faktoren-Analyse untersucht. In diesem Beispiel wirkten die meisten der untersuchten Angebotsmerkmale sowohl als Penalty- wie auch als Reward-Faktoren. Das Preis-Leistungs-Verhältnis stellt dagegen einen reinen Penalty-Faktor dar, während das Merkmal ÖV-Personal zu den Reward-Faktoren zu zählen ist.

2.4.3 Dekompositionelle Verfahren

Während sich bei den bisher besprochenen merkmalsorientierten Verfahren das globale Qualitätsurteil aus den einzelnen Teilurteilen ableitet, wird bei den dekompositionellen Ansätzen der umgekehrte Weg beschritten: Durch die Kombination der Ausprägungen verschiedener Qualitätsmerkmale werden unterschiedliche Entscheidungssituationen (sog. „Vignetten“) nachgebildet (z.B. Pauschalpreis 2 €/ Fahrt; Betriebszeit 5 bis 24 Uhr; 10-Minuten-Takt), die der Befragte einer Nutzenbewertung unterziehen soll (mittels Rating oder Ranking). Möglich ist auch die Abfrage der Verhaltensabsicht. Im Rahmen einer **Conjoint-Analyse** (vgl. hierzu z.B. Backhaus et al. 2006) können dann aus den erhobenen Globalurteilen die partiellen Beiträge der einzelnen Merkmale zum Zustandekommen dieses Gesamturteils berechnet werden. Es geht hierbei somit weniger um die Ermittlung der Zufriedenheit mit einem bestehenden Angebot als vielmehr um die Analyse von Präferenzstrukturen der Kunden. Es soll die Bedeutung einzelner Merkmale

für den Entscheidungsprozess festgestellt werden, um hieraus eine kundenorientierte Produkt- und Preisgestaltung ableiten zu können (vgl. Herrmann, Bauer & Herrmann 1996; Kriemeyer 1990).

Der Vorteil des Verfahrens besteht darin, dass die einzelnen Qualitätskriterien nicht isoliert, sondern gewissermaßen simultan abgefragt werden. Nachteilig ist jedoch, dass bei vielen zu untersuchenden Merkmalen mit vielen Ausprägungen die Zahl der vom Befragten zu beurteilenden Vignetten schnell in astronomische Größenordnungen gerät⁴.

Praxisbeispiele:

Im November und Dezember 1995 wurde in sechs deutschen Städten eine empirische Untersuchung zum Schienennahverkehr durchgeführt. Da eine Methode benötigt wurde, die die Nutzensvorteile der Fahrgäste hinsichtlich eines bestimmten Angebotes offen legt, wurde das Conjoint-Measurement gewählt. Von insgesamt 180 Personen wurden die Präferenzen bezüglich der Gestaltung des Schienennahverkehrs erfasst. Eine Vorstudie ergab, dass aus Sicht der Kunden die sechs Merkmale Betriebszeit, Ticketpreis, Ausstattung der Wagen, Taktzeit während des Hauptverkehrs, Pkw-Parkgelegenheiten an den Haltestellen in den Außenbezirken und die Anschlussmöglichkeit an Busse in die Außenbezirke relevant sind. Aus diesen Merkmalsausprägungen lassen sich 576 unterschiedliche Transportleistungen konstruieren. Um den Aufwand zu beschränken und die Testpersonen nicht zu überfordern, wurde ein experimentelles Design verwendet, das eine Reduzierung auf 16 tatsächlich existierende Angebote ohne einen wesentlichen Informationsverlust erlaubt. Die Testpersonen sollten im nächsten Schritt die 16 Transportleistungen hinsichtlich ihrer Präferenz ordnen. Die hieraus entstehende Rangordnung wurde als Input für das Conjoint Measurement verwendet. Die Auswertung dieser Daten führte zu Teilpräferenzfunktionen für die sechs Merkmale. Die Werte lassen sich nun über ihre Eigenschaften miteinander vergleichen, so dass die relative Bedeutung der Nutzungsbereiche angezeigt werden kann. Aus den Ergebnissen ging hervor, dass der Preis mit 37,0 %, gefolgt von der Betriebszeit mit 20,6 % und der Taktzeit mit 19,9 % das wichtigste Merkmal bildet. Die einzelnen Größen der Preisfunktion bieten eine Möglichkeit, den Wert einzelner Leistungskomponenten zu monetarisieren (vgl. Herrmann, Bauer & Herrmann 1996).

In der Region Hannover wurde das Marktpotenzial verschiedener Mobilitätspakete mit einer Conjoint-Analyse untersucht. Im Rahmen einer mündlich-persönlichen Befragung wurden re-

⁴ Beispielsweise würden sich bei 5 Merkmalen mit jeweils 3 Ausprägungen in einem vollständigen Design 243 Vignetten ergeben.

regelmäßige ÖPNV-Nutzer zu insgesamt 18 verschiedenen Paketalternativen befragt und gebeten, diese in eine Präferenzreihenfolge zu bringen (vgl. Mattern, Zielstorff & Raabe 2001).

Weitere Anwendungsbeispiele:

- Burgenlandbahn: Conjoint-Analyse zum Kundennutzen (Schulz & Meinhold 2003)
- Messung des Kundennutzens mittels Choice-Based-Conjoint-Analyse (Schulz & Meinhold 2003)
- Alternative Bedienformen im ÖPNV – Akzeptanzstudie im Landkreis Saalkreis (Vetter & Haase 2008)

Anmerkung

Von entscheidender Bedeutung für alle merkmalsorientierten Verfahren ist natürlich die adäquate Auswahl der zu untersuchenden Qualitätseigenschaften. Das Auffinden der für die Kunden relevanten Merkmale sollte daher auf der Basis entsprechender Voruntersuchungen erfolgen. Hierfür kommen z.B. Gruppendiskussionen, aber auch die oben angesprochenen ereignisorientierten Verfahren in Frage.

2.5 Kundenzufriedenheitsmessung mittels Befragung: Gruppendiskussionen

Befragungsmethoden lassen sich u.a. danach unterscheiden, ob Interviews mit Einzelpersonen oder Personengruppen geführt werden. Eine verbreitete Form von Gruppeninterviews sind Gruppendiskussionen, unter die man auch so genannte Kundengremien subsumieren kann.

Ganz generell ist die Aufgabe von Gremien, Probleme zu diskutieren sowie Lösungen zu suchen und zu bewerten. Bedingt durch gruppenspezifische Prozesse erhofft man sich von Gremien (z.B. Kundenbeiräte, Fokusgruppen) tiefer gehende Informationen als von Einzelinterviews. Sowohl Kundenbeiräte als auch Fokusgruppen werden auf Initiative des Unternehmens gegründet; sie haben in der Regel ein Vorschlagsrecht, aber keine direkte Entscheidungsbefugnis.

Praxisbeispiel:

Als frühes Beispiel für einen Kundenbeirat kann das 1993 auf Initiative der Bremer Straßenbahn AG gegründete (zeitlich befristete) Fahrgast-Forum genannt werden. Es bestand aus 16 Personen, denen relativ detaillierte Einblicke in betriebliche Abläufe und verkehrliche Konzeptionen ermöglicht wurden (Betriebsbesichtigungen, Gespräche mit Mitarbeitern u.a.). Zentrale Aufgabe

war die Durchleuchtung des ÖPNV-Angebots in Bremen und die Ausarbeitung entsprechender Verbesserungsvorschläge (Pietsch, Lemmermann & Brunjes 1994).

Heute sind Fahrgastbeiräte in zahlreichen Regionen institutionalisiert und werden häufig von den lokalen Akteuren gemeinsam durchgeführt. So gab es 2006 bereits ca. 100 Fahrgastbeiräte in Deutschland (Schiefelbusch & Fliegel 2006).

Fokusgruppen dagegen bestehen nicht ausschließlich aus Kunden, werden von einem Moderator geleitet und stellen ein Forum für Diskussionen und Ideen dar. Hauptzweck ist hier eher ein begleitendes „Monitoring“ der unternehmerischen Aktivitäten mittels laufender und zeitnaher Rückmeldungen und damit die Einbeziehung von Kundenwünschen und -interessen vor allem im Hinblick auf die Produktentwicklung.

In Bezug auf die Messung von Kundenzufriedenheit kann gesagt werden, dass die Sicht der Gremienmitglieder möglicherweise nicht das Qualitätsempfinden der großen Masse der Kunden repräsentiert.

Weitere Anwendungsbeispiele:

- Verkehrsverbund Bremen-Niedersachsen: Fahrgastbeirat als Beitrag zur Kundenorientierung (Herr 1998)
- Großraum Linz: Benutzerbeirat für die Nahverkehrsplanung (Höfler & Czermak 2004)

2.6 Analyse prozessproduzierter Daten

Unter prozessproduzierten Daten versteht man solche, die im normalen betrieblichen Ablauf (z.B. für Zwecke der Pünktlichkeitsstatistik etc.), also ohne den Eingriff von Forschern, entstehen.

So können z.B. dem Betriebsleitsystem (RBL) Informationen zur Pünktlichkeit von Fahrzeugen (z.B. Verspätung in Minuten oder Sekunden) oder auch zum Erreichen von Anschlüssen entnommen werden. Der VDV schlägt eine solche Vorgehensweise unter dem Stichwort „Direkte Leistungsmessung“ vor (VDV 2002, S. 32 ff). Der Vorteil solcher „nicht-reaktiver“ Verfahren ist, dass der Messvorgang das Ergebnis nicht beeinflussen oder verfälschen kann. Solche Daten eignen sich daher insbesondere für den Vergleich von angestrebter und erbrachter Dienstleistungsqualität im Rahmen von Bonus-/ Malus-Vereinbarungen. Ob die angestrebte jedoch auch die vom Kunden erwünschte Qualität abbildet, ist eine andere Frage.

Erwähnt sei noch, dass die direkte Leistungsmessung über die Verwendung rein prozessproduzierter Daten hinausgeht, da in diesem Zusammenhang durchaus auch spezielle Daten erhoben werden (z.B. zum Zustand von Haltestellen und Fahrzeugen), die nicht unbedingt aus dem normalen Betriebsablauf hervorgehen (VDV 2002; vgl. auch DIN EN 13816).

3 Durchführung von Kundenzufriedenheitserhebungen

3.1 Methodische Hinweise zur Durchführung von Kundenzufriedenheitsbefragungen

Im Folgenden sollen zunächst einige allgemeine methodische Hinweise zum Ablauf und zur Durchführung von Kundenzufriedenheitsbefragungen gegeben werden (vgl. hierzu auch Schnippe 1998).

3.1.1 Festlegung der Untersuchungsziele

Der erste Schritt einer Kundenzufriedenheitsbefragung sollte darin bestehen, sich die Ziele und den Zweck der Befragung klar zu machen. So kann es z.B. ganz allgemein um Informationen zum generellen Zufriedenheitsniveau oder um spezielle Themen wie z.B. die Bewertung des Nachtbusangebots gehen. Auch die Evaluation von Angebotsmaßnahmen oder die Ermittlung von Boni bzw. Mali können Gegenstand einer Kundenzufriedenheitsbefragung sein. Im letzteren Fall müssen konkrete Bewertungskriterien festgelegt werden (wann kann von Zufriedenheit bzw. Unzufriedenheit gesprochen werden?). Schließlich müssen im Rahmen der Zielexplication auch noch die relevanten Zielgruppen der Befragung definiert werden (z.B. Abo-Kunden, Fahrgäste einer bestimmten Linie etc.).

3.1.2 Befragungsdesign

Vor der konkreten Planung der Erhebung, also der Festlegung des Befragungsdesigns, sollte ein vorläufiger Arbeits-, Zeit- und Kostenplan erstellt werden. Die Entwicklung des Befragungsdesigns selbst besteht im Wesentlichen aus den Elementen

- Stichprobenauswahl
- Auswahl der Messmethode
- Festlegung der Befragungsform
- Festlegung der Untersuchungshäufigkeit und -regelmäßigkeit,

die im Folgenden erläutert werden.

1. Stichprobenauswahl

Im Rahmen der Stichprobenauswahl ist zunächst die *Grundgesamtheit* der Befragung in sachlicher, zeitlicher und räumlicher Hinsicht abzugrenzen. Beispiele hierfür sind:

- a) alle Abo-Kunden im Verkehrsverbund xy zum Stichtag z
- b) alle Einwohner (ab 10 Jahre) im Bedienungsgebiet xy zum Stichtag z
- c) alle Fahrgäste der Linien a, b und c im Zeitraum Oktober/ November

Im nächsten Schritt geht es darum, für die so definierte Grundgesamtheit eine geeignete *Auswahlgrundlage* für die Stichprobenauswahl bereitzustellen. Für die eben genannten Beispiele stehen folgende Auswahlgrundlagen zur Verfügung:

- a) Kundenkartei (bzw. -datei)
- b) Einwohnermelderegister oder Telefon-/ Adressverzeichnisse
- c) keine Auswahlgrundlage vorhanden

Das *Stichprobenauswahlverfahren* hängt entscheidend davon ab, ob eine (geeignete) Auswahlgrundlage vorhanden ist oder nicht. In den Fällen a) und b) handelt es sich um so genannte Bestandsmassen, hier kann eine einfache (uneingeschränkte) Zufallsauswahl von Personen (Einwohnermelderegister) bzw. Haushalten (Telefonverzeichnis) gezogen werden (z.B. über Zufallszahlen). Da in zunehmendem Maße Haushalte mit Telefonanschluss nicht in den öffentlich zugänglichen Telefonverzeichnissen enthalten sind, wird die Stichprobenauswahl inzwischen häufig durch zufällig generierte Telefonnummern (RDD: Random Digit Dialing) vorgenommen. Soll in den über das Telefonbuch gezogenen Haushalten jeweils nur eine Person befragt werden, muss diese über entsprechende Techniken (z.B. Last-Birthday-Methode oder Schwedenschlüssel) zufällig ausgewählt werden. In diesem Fall nimmt die Auswahlwahrscheinlichkeit einer Person mit der Haushaltsgröße ab, was im Rahmen der Gewichtung/ Hochrechnung wieder ausgeglichen werden muss.

Im Fall c) handelt es sich dagegen um eine Ereignismasse, für die keine Auswahlgrundlage existiert. Insofern ist hier ein mehrstufiges Stichprobenverfahren erforderlich. Auf der ersten Stufe sind dabei zunächst Wagenfahrten (Kurse) auszuwählen. Unter der Annahme, dass die Einsatzpläne der betreffenden Fahrzeuge im Vorhinein bekannt und im Idealfall in elektronischer Form verfügbar sind, kann - im Rahmen eines vorher festzulegenden Erhebungszeitraums - die Auswahl der Kurse nach dem Prinzip einer einfachen Zufallsstichprobe stattfinden. Denkbar ist je nach Datenlage auch eine geschichtete Zufallsauswahl, bei der die Gesamtheit

der Kurse vor der Stichprobenauswahl nach Linien und/ oder zeitlichen Merkmalen untergliedert wird. Für jede der so gezogenen Wagenfahrten erfolgt auf der zweiten Stufe die Auswahl der konkreten Befragungspersonen, die im Idealfall als Vollerhebung (einfache Klumpenstichprobe) oder im Regelfall ebenfalls als Zufallsauswahl (zweistufige Stichprobe) durchgeführt werden sollte. Eine einfache Möglichkeit bestünde beispielsweise darin, im Fahrzeug nach jedem Halt an einer Haltestelle die jeweils erste Person, die das Fahrzeug betritt, zu interviewen bzw. für ein Interview zu rekrutieren. Hier entspricht dann die Zahl der pro Kurs zu befragenden Fahrgäste genau der Zahl der Haltestellen (ohne Endhaltestelle), sofern an jeder Haltestelle mindestens eine Person einsteigt und an der Erhebung teilnimmt. Aus Annahmen über die mittlere Haltestellenzahl pro Kurs lässt sich unter Berücksichtigung des Gesamtstichprobenumfangs (Anzahl zu befragender Fahrgäste) die Zahl der auszuwählenden Wagenfahrten bestimmen. Alternativ wäre daran zu denken, im Sinne eines „Schwedenschlüssels“ unter den jeweiligen Einsteigern die zu interviewende Person nach einem Zufallsverfahren festzulegen. Wegen des möglichen Klumpeneffekts (Homogenität der an einer Haltestelle einsteigenden Fahrgäste) erscheint ein Verfahren vorteilhafter, das einer einfachen Zufallsauswahl aus der Gesamtheit der (Erst-)Einsteiger eines Kurses nahekommt. Man könnte dies in Form einer systematischen Zufallsauswahl (in zeitlicher Reihenfolge und unabhängig von der Einstiegshaltestelle wird jeder k-te Einsteiger - z.B. jeder zehnte - interviewt) realisieren. Hier kann die erwartete Zahl der Interviews pro Kurs aus Daten über Linienbelastungen abgeleitet werden. Wenn die ausgewählte Person die Teilnahme verweigert oder nicht zur Zielgesamtheit (z.B. Personen ab 14 Jahre) gehört, wird die nächste Person angesprochen usw.

Da bei einer solchen Vorgehensweise aus der Gesamtheit der Fahrten bzw. Beförderungsvorgänge ausgewählt wird, ist die Auswahlwahrscheinlichkeit einer Person umso höher, je mehr Fahrten sie durchführt. Die resultierende Stichprobe repräsentiert somit die Verteilung der Fahrgäste nach deren Nutzungshäufigkeit. Ganz anders ist dies bei einer Auswahl von Personen (bspw. solche, die in den letzten 12 Monaten mindestens ein Mal öffentliche Verkehrsmittel benutzt haben) z.B. mittels „Screening“ aus einem Adressverzeichnis. Hier ist die Auswahlwahrscheinlichkeit unabhängig von der Nutzungshäufigkeit, dementsprechend werden Selten- und Gelegenheitsnutzer i.d.R. häufiger vertreten sein, als es ihrem Anteil am Fahrtenaufkommen entspricht. Dies ist bei der Interpretation von Ergebnissen aus Kundenzufriedenheitsbefragungen zu berücksichtigen.

Was schließlich die *Festlegung der Stichprobengröße* betrifft, so hängt der erforderliche Mindeststichprobenumfang neben dem Auswahlverfahren vor allem

- von der gewünschten Schätzgenauigkeit (absoluter bzw. relativer Fehler),
- vom Sicherheitsgrad (i.d.R. 95 %) und
- von der Auswahl des Zielmerkmals

ab (vgl. hierzu auch VDV 2010). Um im Rahmen der Stichprobenplanung Mindeststichprobenumfänge berechnen zu können, müssen zahlenmäßige Vorstellungen über die Streuung mindestens eines Untersuchungsmerkmals vorliegen.

Im folgenden Beispiel für das Zielmerkmal Globalzufriedenheit wird vereinfachend zunächst eine einfache Zufallsauswahl unterstellt. Die Vorgabe soll lauten, dass der zu schätzende Mittelwert für die Globalzufriedenheit (5-stufige Skala von 1 bis 5) von der wahren mittleren Globalzufriedenheit in der Grundgesamtheit aller Kunden mit einer Wahrscheinlichkeit von 95 % (Sicherheitsgrad $1-\alpha=0,95$, d.h. $t_{1-\alpha/2}=1,96$) um nicht mehr als 5 % nach unten oder oben abweicht (maximaler relativer Fehler der Schätzung $r=0,05$). Als Grundlage für die nachfolgende Beispielrechnung wird für die Globalzufriedenheit der Mittelwert $m=2,5$ und die Varianz $s^2=0,93$ angesetzt. Dies sind Durchschnittswerte, welche aus früheren Befragungen gewonnen wurden.

Für den erforderlichen Stichprobenumfang der geplanten Befragung erhält man bei den genannten Genauigkeitsanforderungen folgenden Wert:

$$n_0 = (t^2 s^2) / (r^2 m^2) = (3,8416 \cdot 0,93) / (0,0025 \cdot 6,25) = 228,65, \text{ d.h., es sind 229 Kunden zu befragen.}$$

Bei diesem Stichprobenumfang wird der geforderte Genauigkeitsstandard erreicht. Man kann bei diesem Stichprobenumfang also praktisch sicher sein (Sicherheitsgrad 95 %), dass die „wahre“ mittlere Globalzufriedenheit sich von dem in der durchzuführenden Befragung ermittelten Wert um nicht mehr als 5 % unterscheidet. Dies entspricht einem absoluten Fehler von 0,125 „Skalenpunkten“ und einem zugehörigen 95 %-Konfidenzintervall $[2,375; 2,625]^5$. Untergliedert man die Stichprobe z.B. nach Altersgruppen, so sind die gruppenspezifischen Schätzungen natürlich mit einem größeren Stichprobenfehler behaftet.

Da insbesondere bei Fahrgastbefragungen die Auswahl der zu befragenden Personen nach einem mehrstufigen Verfahren und nicht nach dem Verfahren der einfachen Zufallsauswahl

⁵ Sofern die Schätzung für den Mittelwert tatsächlich 2,5 beträgt.

erfolgt (siehe oben), sind in diesem Fall die angegebenen Werte für den absoluten Fehler bei gegebenem Stichprobenumfang und auch die Mindeststichprobenumfänge bei gegebenen Genauigkeitsanforderungen zu niedrig. Zur Berücksichtigung der Mehrstufigkeit sollte der den obigen Berechnungen zugrunde liegende Standardfehler der Schätzung mit einem Faktor größer als Eins (in der Stichprobentheorie als Design-Effekt bezeichnet) multipliziert werden. Die genaue Bestimmung dieses Faktors ist schwierig und nur auf der Grundlage von Daten aus früheren Fahrgastbefragungen zu vergleichbaren Themen möglich. Ohne weitergehende Informationen sollte man den Standardfehler gemäß einfacher Zufallsauswahl mit dem Faktor 1,5 (ggf. auch Faktor 2) multiplizieren, um den Stichprobenumfang nicht zu klein zu wählen bzw. die tatsächliche Hochrechnungsgenauigkeit nicht zu optimistisch einzuschätzen.

2. Auswahl der Messmethode

Bei den auf der Befragung von Kunden basierenden Methoden zur Messung der Kundenzufriedenheit ist zwischen ereignisorientierten und merkmalsorientierten Verfahren zu differenzieren. Die entsprechenden Methoden wurden in den Abschnitten 2.3 und 2.4 im Überblick vorgestellt. Die multiattributive Messung (merkmalsorientiert) und die Critical Incident-Technik (ereignisorientiert) werden in den Kapiteln 3.3 und 3.4 unter Berücksichtigung von Ergebnissen des dem Leitfadens zugrunde liegenden Forschungsprojektes „Bedeutung von Kundenzufriedenheit im ÖPNV“ nochmals ausführlich behandelt.

3. Festlegung der Befragungsform

Bei der Befragungsform kann man zwischen schriftlichen (schriftlich-postalischen), mündlich-persönlichen, telefonischen und Online-Befragungen (per E-Mail oder Internet) unterscheiden. Grundsätzlich ist die Befragungsform von der benutzten Auswahlgrundlage unabhängig, allerdings müssen in einigen Fällen zusätzliche Informationen beschafft werden. Wird z.B. im Rahmen einer Haushaltsbefragung⁶ die Stichprobe aus der Liste der Abo-Kunden oder dem Einwohnermelderegister gezogen, so müssen in aller Regel für eine telefonische Befragung die Telefonnummern recherchiert werden⁷. Soll die Erhebung online stattfinden, kann z.B. die URL eines im Internet hinterlegten Fragebogens per Post oder, falls bekannt, per E-Mail zugesandt werden. Basiert die Stichprobe dagegen auf einer Auswahl aus dem Telefonbuch, so fehlen hier

⁶ Bei einer Haushaltsbefragung findet das Interview in der Wohnung der zu befragenden Person statt.

⁷ Sollen auf Basis des Einwohnermelderegisters nur Kunden telefonisch befragt werden, ist ein entsprechendes „Screening“ notwendig, d.h., es muss zu Beginn des Interviews geklärt werden, ob die Person zur anvisierten Grundgesamtheit gehört. Dies gilt natürlich auch für eine telefonische Befragung auf Basis einer Telefonbuchstichprobe.

vielfach die Adressangaben, was andere Erhebungsformen als die telefonische Befragung erschwert.

Auch bei Fahrgastbefragungen können prinzipiell alle Befragungsformen angewandt werden. Bei der telefonischen Variante werden Kunden im Fahrzeug (oder an der Haltestelle) angesprochen und gefragt, ob sie bereit wären, an einer Befragung teilzunehmen. Falls ja, wird die Telefonnummer aufgenommen und im Verlauf der nächsten Tage⁸ ein telefonisches Interview geführt. Die Kontaktdaten können z.B. mittels Handy/ PDA aufgenommen werden, mit denen bereits vor Ort einige grundlegende Informationen abgefragt und erfasst werden können und die eine direkte Übertragung dieser Erstkontaktdaten vom Interviewer an das Telefonstudio und somit sehr zeitnahe Telefoninterviews ermöglichen.

Bei einer schriftlich-postalischen Befragung „durch Verteilen“ werden Fragebögen (mit Rückumschlag) in Fahrzeugen oder an Haltestellen den ausgewählten Personen persönlich übergeben, wobei diese den Bogen nach dem Ausfüllen zu Hause per Post zurücksenden sollen⁹. Bei dieser Variante empfiehlt es sich, die Fragen nicht auf den konkreten Nutzungsvorgang zu beziehen, da der Zeitraum zwischen Übergabe und Beantwortung des Fragebogens von Person zu Person unterschiedlich lang ist und somit häufig nicht von einer korrekten Erinnerung an die entsprechende Kontaktsituation ausgegangen werden kann. Da eine Befragung nach einem solchen Design absolut anonym abläuft, kann keine Rücklaufkontrolle oder Nachfassaktion durchgeführt werden. Daher sind die dabei erzielbaren Rücklaufquoten in der Regel eher gering. Bei einer rein schriftlichen Befragung würde man die ausgegebenen Fragebögen zeitnah bzw. an Ort und Stelle ausfüllen lassen und später wieder einsammeln (vgl. z.B. Faltlhauser 1999).

Bei mündlich-persönlichen Interviews vor Ort wird die Befragung durch geschulte Interviewer im Fahrzeug oder an der Haltestelle mittels Papierfragebögen oder computerunterstützt durchgeführt. Da die Aufenthaltsdauer des Kunden im Fahrzeug oder an der Haltestelle sehr kurz sein kann, sollten hier keine allzu langen Fragebögen eingesetzt werden. Alternativ wäre natürlich auch die Erfassung der Adresse mit anschließender Vereinbarung eines Interviewtermins in der Wohnung des Befragten denkbar.

⁸ Wenn es um die Bewertung des konkreten Nutzungsvorgangs geht, sollte das Interview möglichst zeitnah, im Idealfall noch am Abend desselben Tages stattfinden.

⁹ Denkbar wäre natürlich auch die Erfassung der Post- bzw. E-Mail-Adresse mit anschließender Zustellung eines entsprechenden (Online-)Fragebogens.

Generell ist bei Fahrgasterhebungen zu bedenken, dass damit nur aktuelle Kunden erreicht werden können. Personen, die z.B. wegen negativer Erfahrungen öffentliche Verkehrsmittel nicht mehr nutzen, können mit dieser Methode naturgemäß nicht befragt werden.

Im Zusammenhang mit der Auswahl der Befragungsform ist auch über den *Grad der Standardisierung* der Fragen zu entscheiden. Hier reicht die Spannweite von voll standardisierten Fragebögen über leitfadengestützte Interviews bis hin zu so genannten „narrativen Interviews“, also dem völlig freien Erzählen des Befragten. Die zu wählende Befragungsform ist insofern nicht völlig unabhängig vom Messverfahren, als ereignisorientierte Verfahren eher qualitativ ausgerichtet sind, so dass hier insbesondere leitfadengestützte mündlich-persönliche Interviews in Betracht gezogen werden sollten.

Auf Erhebungsinhalte und die Formulierung von Fragen und Antwortvorgaben wird in den Kapiteln 3.3 und 3.4 eingegangen. Standardmäßig sollten am Ende der Befragung immer soziodemographische Merkmale (Alter, Geschlecht, Pkw-Verfügbarkeit) sowie Angaben zum Nutzerprofil (ÖPNV-Nutzungshäufigkeit, bei Fahrgastbefragungen auch Fahrtzweck und benutzter Fahrchein) erhoben werden.

4. Festlegung der Untersuchungshäufigkeit und -regelmäßigkeit

Bei der Planung einer Kundenzufriedenheitsbefragung ist letztlich noch zu entscheiden, ob die Erhebung einmalig stattfinden (Querschnittsbefragung) oder - ggf. auch mehrfach - wiederholt werden soll (Längsschnittbefragung). Dies hängt natürlich entscheidend von den Untersuchungszielen ab. Soll z.B. ein „laufendes Zufriedenheitsmonitoring“ aufgebaut werden oder ist eine Maßnahmenevaluation angestrebt, so ist eine Längsschnittuntersuchung angezeigt. Hier ist dann zum einen die Frage zu beantworten, in welchen zeitlichen Abständen die einzelnen Erhebungswellen stattfinden sollen, zum anderen ist zu klären, ob von Welle zu Welle eine neue Stichprobe gezogen oder ein Panel-Design angewandt werden soll. Im Rahmen eines Panels werden dieselben Personen mit demselben Erhebungsinstrument mindestens zweimal befragt. Generell liegen die Vorzüge eines Panel-Designs darin, dass die Schätzung der Veränderung zwischen zwei Zeitpunkten genauer (als die aus zwei unabhängigen Stichproben resultierende Schätzung der Veränderung) ist, falls das Untersuchungsmerkmal über die Untersuchungszeitpunkte korreliert¹⁰.

¹⁰ Die Streuung einer Differenz aus zwei Zeitpunkten bei verbundenen Stichproben ist ja die Summe der beiden Einzelvarianzen abzüglich zweimal der Kovarianz (vgl. z.B. Hartung 2009).

Es bietet sich an, Längsschnitterhebungen turnusmäßig (jedes Quartal oder Jahr) durchzuführen oder zu bestimmten inhaltlich begründeten Zeitpunkten, z.B. wenn sich grundlegende Angebotsparameter geändert haben. Der Einfluss externer temporärer Rahmenbedingungen (z.B. 6-monatige Großbaustelle oder gewaltsamer Übergriff im Bereich der U-Bahn), kann bei Längsschnitterhebungen besser identifiziert und quantifiziert werden als bei einmalig durchgeführten Erhebungen.

3.1.3 Vorbereitung und Durchführung der Befragung

Unabhängig von der Befragungstechnik sollte zumindest bei neuen oder stark modifizierten Fragebögen ein Pretest durchgeführt werden. Der Verzicht auf Tests zur Überprüfung der Leistungsfähigkeit eines Fragebogens vor Beginn der Feldarbeit erhöht die Fehlerwahrscheinlichkeit bei der Umsetzung von Kundenzufriedenheitsbefragungen. Fehler aufgrund der Fehlinterpretation von Fragen, dem Fehlen von Informationen und falschen Antworten von Befragten können mit Hilfe eines Pretests vermieden werden.

Wesentliche Voraussetzung für verlässliche Ergebnisse bei mündlich-persönlichen Befragungen ist ein Pool an zuverlässigen und einsatzbereiten Interviewern mit sicherem Auftreten in der Öffentlichkeit. Kurz vor Erhebungsbeginn sollten die Interviewer im Rahmen einer Schulung in ihre Aufgaben eingewiesen werden. Bei der Schulung sollten neben dem Zweck, dem Inhalt und dem Ablauf der Befragung organisatorische Fragen behandelt werden (z.B. Dienstplan, Einsatzplan, Personalabrechnung, bei Papierfragebögen Fragebogenlogistik) und Kenntnisse im Zusammenhang mit dem Erhebungsgegenstand (z.B. Tarifbestimmungen im Untersuchungsraum) sowie - bei computergestützten Verfahren - zum Umgang mit der Hard- und Software vermittelt werden.

Im Folgenden werden noch einige allgemeine Hinweise zur Durchführung von Befragungen in der Differenzierung nach den einzelnen Befragungsformen gegeben (vgl. z.B. Bäumler et al. 2010, Porst 2000; Dillman 1999; Richardson et al. 1995):

Mündlich-persönliche Interviews

Ein Problembereich bei mündlich-persönlichen Interviews sind Interviewer-Effekte und Interviewer-Fehler, so dass der Auswahl, Schulung, Betreuung und Überwachung des Erhebungspersonals eine hohe Bedeutung zukommt. Durch den persönlichen Kontakt treten bei dieser Kommunikationsart allerdings meist geringere Antwortausfälle auf als bei telefonischen oder schriftlich-postalischen Befragungen. Darüber hinaus kann der Interviewer auf die Befragten näher eingehen, Verständnisprobleme klären und durch Nachfragen genauere und eindeutige Infor-

mationen erlangen, als dies bei schriftlichen Befragungen möglich wäre. Persönliche Interviews eignen sich - von Befragungen im Fahrzeug abgesehen - daher besonders, um schwierige und komplexe Sachverhalte zu erheben.

Schriftlich-postalische Befragungen

Eine Klärung von Verständnisproblemen ist bei schriftlichen Befragungen nicht möglich, da im Gegensatz zu persönlichen Interviews der direkte Kontakt zwischen Interviewer und Befragtem fehlt. Mängel des Fragebogens führen daher bei schriftlichen Befragungen zu deutlich mehr Fehlern als bei persönlichen Interviews. Im Vergleich zu den mündlichen Verfahren kommt der Gestaltung des Fragebogens eine wesentlich höhere Bedeutung zu. Darüber hinaus ist die Interviewsituation nicht kontrollierbar, d.h. es kann nicht gewährleistet werden, dass die zu interviewende Person den Fragebogen auch tatsächlich selbst ausgefüllt hat.

Einen Einfluss auf die Qualität der Ergebnisse hat auch die äußere Form des Fragebogens. Dabei sollten Aspekte wie lesbare Schrift, übersichtliche Anordnung der Fragen oder Gliederung des Bogens durch farbliche Strukturierung berücksichtigt werden. Erläuterungen zum Ausfüllen der Fragebögen in Form von verbalen Ausfüllhinweisen – ggf. auch in Form eines Ausfüllbeispiels – führen zu einer deutlichen Verbesserung der Qualität.

Folgender Ablauf hat sich bei einem postalischen Versand bewährt:

1. Versand eines Ankündigungsschreibens ca. 14 Tage vor dem Erhebungsstichtag.
2. Hauptversand der Befragungsunterlagen ca. vier Tage vor dem Erhebungsstichtag; zu den Befragungsunterlagen gehören
 - ein (ggf. von amtlicher Stelle) unterzeichnetes Motivationsschreiben (Inhalt: Zweck der Erhebung, Notwendigkeit der Beteiligung, Zusicherung der Anonymität, grundsätzliche Ausfüllhinweise),
 - eine Erklärung zum Datenschutz durch das durchführende Institut,
 - die Fragebögen,
 - ein Rücksendekuvert.
3. Versand einer Erinnerungspostkarte ca. eine Woche nach dem Erhebungsstichtag.

Neben dem postalischen Versand der Unterlagen ist auch die persönliche Übergabe der Unterlagen möglich. Der persönliche Kontakt ist zwar i. d. R. aufwändiger als der Postversand, hat aber andererseits positive Effekte hinsichtlich der Ausschöpfung.

Telefonische Interviews (CATI - Computer Assisted Telephone Interviewing)

Zur Vermeidung systematischer Fehler sollten folgende organisatorische Maßnahmen bei einer telefonischen Befragung umgesetzt werden:

- eine möglichst umfassende Betreuung der ausgewählten Haushalte z.B. über Ankündigungsschreiben und/oder eine Flankierung über Pressemitteilungen, um die Teilnahmebereitschaft zu erhöhen
- mehrere Versuche zur Kontaktaufnahme, teilweise auch zu unterschiedlichen Zeiten;
- die Möglichkeit, einen Termin für die Befragung zu vereinbaren.

Online-Befragungen (CASI - Computer Assisted Self Interviewing)

Zunehmend an Bedeutung gewinnen Online-Befragungen, da diese erhebliche Zeit- und Kostenvorteile bieten. Aus methodischer Sicht ist zu empfehlen, dass die Ziehung der Stichprobe wie bei den „klassischen“ Verfahren erfolgt (z.B. aus einem amtlichen Einwohnermelderegister). Anschließend wird z.B. durch eine kurze schriftliche oder telefonische Anfrage die E-Mail-Adresse ermittelt und dann der zu befragenden Person eine Verlinkung auf einen Online-Fragebogen zugeschickt. Das weitere Vorgehen orientiert sich an dem einer schriftlichen Befragung, wobei die Befragung hier über ein anderes Medium, das Internet, durchgeführt wird und ggf. vorgesehene Erinnerungsschreiben dann natürlich auch per E-Mail versandt werden. Ein weiterer Vorteil dieser Befragungsart ist, dass jüngere Zielgruppen einfacher für eine Befragung gewonnen werden können, zumal solche Fragebögen z.B. auch über Smartphones oder Tablet-PCs ausgefüllt werden können. Inzwischen wird auch über so genannte „Online-Panel“ erhoben, d.h. es gibt einen Pool von auf verschiedene Art und Weise (teilweise Selbstauswahl) rekrutierten Personen, von denen die E-Mail-Adresse bekannt ist und welche schnell für eine Befragung bereit stehen. Dies hat natürlich praktische Vorteile, aus methodischer Sicht ist allerdings zu beachten, dass die Teilnahme starken Selbstselektionseffekten (Alter, Technikaffinität) unterliegt. Außerdem sind solche Online-Panels für regionale Befragungen durch die aus statistischer Sicht zu geringe Anzahl von Probanden aus der betreffenden Region zumeist nicht geeignet. Zudem ist noch nicht eindeutig geklärt, ob durch die häufigere Ansprache der Befragten Verzerrungen im Antwortverhalten hervorgerufen werden. Andere Verfahren wie das Ausfüllen und Rücksenden von angehängten PDF- oder Word-Dateien werden nur noch selten praktiziert, da sie für den Befragten zeitlich aufwändiger und oftmals technisch zu kompliziert sind.

Der derzeit gebräuchlichste Einsatz von Online-Befragungen im Verkehrsbereich ist die Kombination mit einer schriftlichen oder telefonischen Befragung (Methodenmix).

Rolle von Incentives

Durch den Einsatz von Incentives bei Befragungen, insbesondere in schriftlich-postalischer Form, lässt sich i. d. R. die Ausschöpfungsquote erhöhen. Besonders bei Panelbefragungen ist je nach Länge und Komplexität der Befragung der Gebrauch von Incentives zu empfehlen. Neben monetären Anreizen besteht die Möglichkeit, nicht-monetäre Incentives wie z.B. Lotterielose, Briefmarken oder Spendenbeiträge einzusetzen. Am effektivsten, d. h. mit der größten Rücklauf steigernden Wirkung, ist der Einsatz von monetären Anreizen. Dabei sollte in keinem Fall der materielle Anreiz zu hoch angesetzt werden, da ansonsten das Incentive aus Sicht der Befragten den symbolischen Charakter eines „kleinen Dankeschöns“ verliert.

Um Verzerrungen der Stichprobenszusammensetzung zu vermeiden, ist unbedingt darauf zu achten, dass Incentives nicht zum zentralen Motiv für die Teilnahme an der Befragung werden und für alle Befragten annähernd gleich relevant sind. Nicht-monetäre Incentives sollten unabhängig vom Untersuchungsgegenstand gewählt werden, damit Teilnehmer durch den Anreiz nicht dazu verleitet werden, von ihrer tatsächlichen Meinung abweichende Antworten zu geben.

3.1.4 Datenaufbereitung und Auswertung

Im Falle einer schriftlichen Befragung sollten die ausgefüllten Fragebögen nach dem Eingang auf Plausibilität geprüft und ggf. bereinigt werden. Anschließend erfolgt die Erfassung der Daten und die Erstellung der Auswertungsdatei. Im Rahmen von computerunterstützten Befragungen erfolgt die Datenerfassung bereits während der Interviews. In jedem Fall sollte die Auswertungsdatei vor Beginn der Datenanalyse auf Erfassungsfehler und Inkonsistenzen in den Daten geprüft werden.

Zur Korrektur von Verzerrungen durch Antwortausfälle (Non-Response) und das Auswahlverfahren kann eine *Gewichtung* erforderlich sein. Bei Stichprobenverfahren mit ungleichen Inklusionswahrscheinlichkeiten (z.B. bei Auswahl je einer Person aus einem Haushalt) wird das Gewicht proportional zum Kehrwert der Auswahlwahrscheinlichkeit eines Elementes im Stichprobenplan gewählt („Transformation“). Zum Ausgleich von durch Non-Response bedingten Verzerrungen wird das Gewicht mittels externer Häufigkeitsverteilungen bestimmt („Redressment“). Hierzu wird eine Quelle gewählt, welche die Grundgesamtheit bezüglich der zur Gewichtung herangezogenen Variablen verlässlich widerspiegelt. Das Gewichtungsverfahren hängt dabei wesentlich davon ab, welche Strukturinformationen über die Gesamtheit verfügbar sind. Im Fall der Rekrutierung von Befragungspersonen im Fahrzeug oder an Haltestellen sollten für jede kontaktierte Person einige Grundinformationen (Geschlecht, Altersgruppe,...) erhoben werden,

um die entsprechenden Verteilungen in der Nettostichprobe mit denen aller kontaktierten Fahrgäste (Bruttostichprobe) vergleichen zu können.

Auf die *Auswertung* von Kundenzufriedenheitsbefragungen wird in den Abschnitten 3.3 und 3.4 eingegangen. Auf der Basis der Ergebnisse kann dann der letzte Schritt einer Kundenzufriedenheitsbefragung, die Ableitung von Planungs- und Entscheidungshilfen für das Management erfolgen.

3.2 Beschreibung der zugrunde liegenden Erhebung

Im Rahmen des zugrunde liegenden Forschungsprojektes wurde eine telefonische Kundenzufriedenheitsbefragung mit 2 Erhebungswellen (Panel) in den Städten Bonn, Ludwigshafen und Tübingen durchgeführt. Diese drei Städte wurden deshalb ausgewählt, weil dort im Zeitraum zwischen den beiden Befragungswellen (1 Jahr) Maßnahmen in Bezug auf das ÖPNV-Angebot eingeführt wurden. In der 1. Erhebungswelle wurden insgesamt 3.153 Personen von der Omniphon GmbH befragt, von denen 1.702 nach etwa einem Jahr erneut interviewt werden konnten.

Pro Gebiet wurden insbesondere aus methodischen Gründen drei Teilstichproben (TSP) unterschieden:

- Bei Teilstichprobe 1 (TSP1) handelt es sich um Personen, die am Befragungstag bzw. maximal zwei Tage vor dem Befragungstag einen Weg mit dem ÖPNV im jeweiligen Gebiet zurückgelegt haben. Die Fragen zur Kundenzufriedenheit beziehen sich hier dann auf einen konkreten ÖPNV-Weg.
- Teilstichprobe 2 (TSP2) besteht aus ÖPNV-Nutzern, die zwar in den letzten drei Tagen nicht mit öffentlichen Verkehrsmitteln unterwegs waren, diese aber zumindest einmal in den letzten 12 Monaten (im jeweiligen Untersuchungsgebiet) genutzt haben.
- Teilstichprobe 3 (TSP3) schließlich sind die Nicht-Nutzer, also diejenigen, die in den letzten 12 Monaten keine Fahrt mit dem ÖPNV im jeweiligen Gebiet gemacht haben.

Als Verfahren zur Messung der Kundenzufriedenheit wurden multiattributive Verfahren sowie die Critical Incident-Technik angewandt. Diese Messverfahren werden im Folgenden ausführlich dargestellt, wobei die methodischen Erläuterungen durch ausgewählte Ergebnisse der durchgeführten Befragung (1. Erhebungswelle) ergänzt werden.

3.3 Multiattributive Verfahren

Nach Hentschel (1992, 2000) lassen sich multiattributive Verfahren anhand dreier Begriffspaare klassifizieren:

- Einstellungs- vs. zufriedenheitsorientierte Messung
- Direkte vs. indirekte Messung
- Ein- vs. Zweikomponentenansatz.

Da sich alle Kriterien miteinander kombinieren lassen, ergeben sich insgesamt 8 Grundvarianten multiattributiver Messung.

3.3.1 Einstellungs- vs. zufriedenheitsorientierte Messung

Einstellungs- und zufriedenheitsorientierte multiattributive Messung unterscheiden sich im Wesentlichen dadurch, dass sich bei der zufriedenheitsorientierten Variante die Einschätzung der einzelnen Qualitätsmerkmale auf die konkrete Nutzung der Dienstleistung bezieht. Dabei wird oft der Fokus auf die letzte Begegnung (letzte ÖPNV-Fahrt) zwischen Kunde und Anbieter gerichtet. Bei der einstellungsorientierten Variante steht dagegen mehr die generelle Bewertung der Qualität des Anbieters oder der Leistung im Vordergrund. Hier können zwar ganz analog auch Zufriedenheiten mit einzelnen Qualitätseigenschaften ermittelt werden, diese sind jedoch nicht auf bestimmte Situationen gerichtet.

Bei der Anwendung dieser Methode muss somit die Festlegung getroffen werden, ob die Zufriedenheit mit einzelnen Qualitätsaspekten allgemein oder auf konkrete Fahrten bezogen abgefragt wird¹¹. Im Rahmen der Kundenzufriedenheitsbefragung wurden beide Varianten getestet. In Teilstichprobe 1 wurde die Zufriedenheit bei der letzten ÖPNV-Fahrt ermittelt, in Teilstichprobe 2 wurde dagegen allgemein nach der Zufriedenheit gefragt. In der folgenden Tabelle sind die Mittelwerte der insgesamt 11 abgefragten Qualitätseigenschaften (Zufriedenheit und Wichtigkeit) für diejenigen Personen aus den Teilstichproben 1 und 2 ausgewiesen, die den öffentlichen Nahverkehr an einem bis 3 Tagen pro Woche nutzen und somit eine ähnliche und vergleichbare Nutzungshäufigkeit aufweisen:

¹¹ Im Prinzip können auch Nicht-Kunden nach ihrer allgemeinen Einstellung gegenüber einzelnen Qualitätseigenschaften gefragt werden, da es für Verkehrsunternehmen durchaus von Interesse sein kann, wie die Angebotsqualität von Nicht-Nutzern eingeschätzt wird, um so Hemmnisse für die ÖV-Nutzung entdecken und beseitigen zu können. Hier ist allerdings u.U. mit einer hohen Zahl an Antwortverweigerungen zu rechnen. In jedem Fall sollte dabei nicht mit dem Begriff „Zufriedenheit“ operiert werden.

Tabelle 1: Wichtigkeit von und Zufriedenheit mit einzelnen Merkmalen des ÖPNV bei regelmäßigen ÖPNV-Nutzern (Mittelwerte)

Merkmal	TSP 1		TSP 2	
	Wichtigkeit (abgefragt)	fahrtbez. Zufriedenheit	Wichtigkeit (abgefragt)	allgem. Zufriedenheit
Anzahl Fahrten/ Verbindungen (Tag)	1,64	1,99	1,84	1,95
Anzahl Fahrten/ Verbindungen (We, Abend)	2,00	2,88	2,04	3,12
Sicherheit der Anschlüsse	1,58	2,40	1,63	2,50
Pünktlichkeit	1,46	2,28	1,55	2,31
Schnelligkeit	1,87	2,01	1,92	2,16
Übersichtlichkeit Liniennetz	1,97	2,14	2,02	2,25
Platzangebot im Fahrzeug	2,03	2,31	2,07	2,23
Information an Haltestellen	1,84	2,28	1,92	2,49
Information im Fahrzeug	2,31	2,27	2,33	2,52
Preis-Leistungsverhältnis	1,61	2,66	1,70	3,00
persönliche Sicherheit an Haltestellen	1,77	2,12	1,77	2,35

Basis: Personen, die an 1 bis 3 Tagen pro Woche öffentliche Verkehrsmittel benutzen (TSP 1, n=348 Befragte; TSP 2, n=311 Befragte)

1 = „sehr zufrieden“ / „sehr wichtig“; 5 = „gar nicht zufrieden“ / „gar nicht wichtig“

Die Ergebnisse zeigen, dass bei Personen mit etwa gleicher ÖPNV-Nutzungshäufigkeit in den beiden Gruppen die fahrtspezifische Zufriedenheit (TSP 1) bei fast allen Merkmalen (einschließlich Preis-Leistungsverhältnis) höher ist als die allgemein abgefragte Zufriedenheit in TSP 2. Dies dürfte damit zu tun haben, dass in das allgemeine Urteil in höherem Maße in der Vergangenheit gemachte (negative) Erfahrungen sowie auch das wahrgenommene „Image“ (vermittelt durch Bekannte, Medien etc.) einfließen. Allerdings ist die Streuung bei den fahrtbezogenen Zufriedenheitseinschätzungen deutlich höher als bei den allgemeinen, d.h. dass Zufriedenheitseinschätzungen, die auf konkrete Fahrten bezogen sind, stärker oszillieren als die allgemeinen Einschätzungen, bei denen die Bewertung über viele Nutzungsvorgänge „gemittelt“ wird. Dies kann bei der fahrtbezogenen Vorgehensweise dazu führen, dass sich besondere Ereignisse im ÖPNV während der Datenerhebung (z.B. wetterbedingte Störungen im Betriebsablauf) in den Ergebnissen niederschlagen.

Fragen und Frageformulierung

Zur Erhebung der Qualitätswahrnehmung stehen im Wesentlichen drei Alternativen zur Wahl (vgl. Hentschel 1992, S. 146):

- Gut/ Schlecht- bzw. Zufrieden/ Unzufrieden-Urteil bei jedem Merkmal
(Bsp.: „Wie zufrieden sind Sie mit der Pünktlichkeit?“ (sehr zufrieden,..., überhaupt nicht zufrieden)
- Erhebung der wahrgenommenen Eigenschaftsausprägung bei jedem Merkmal
(Bsp.: „Wie pünktlich sind (Ihrer Meinung nach) die öffentlichen Verkehrsmittel?“ (sehr pünktlich,..., überhaupt nicht pünktlich)
- Erhebung von Zustimmung- bzw. Ablehnungsgraden zu merkmalsbezogenen Statements
(Bsp.: „Die öffentlichen Verkehrsmittel sind pünktlich“ (stimme voll und ganz zu,..., stimme überhaupt nicht zu)

In der vorliegenden Befragung wurde die erste Variante gewählt:

„Ich nenne Ihnen jetzt einige Eigenschaften des öffentlichen Nahverkehrs. Bitte sagen Sie mir auf einer Skala von 1 bis 5, wie zufrieden Sie damit in <Gebiet> sind.“ (Liste) (einstellungsorientiert) bzw.

„Ich nenne Ihnen jetzt einige Eigenschaften des öffentlichen Nahverkehrs. Bitte sagen Sie mir auf einer Skala von 1 bis 5, wie zufrieden Sie damit bei dieser Fahrt mit <benutztes Verkehrsmittel> waren.“ (Liste) (zufriedenheitsorientiert).

Unabhängig von der gewählten Formulierung ist es dabei wichtig, auf Eindimensionalität zu achten, d.h. es sollte pro Item nur nach einem Qualitätsaspekt gefragt werden. So wäre zum Beispiel die Frage „Wie zufrieden sind Sie mit der Sauberkeit an Haltestellen und in Fahrzeugen?“ in methodischer Hinsicht problematisch.

Generell ist es bei multiattributiven Verfahren natürlich so, dass man nur Informationen über die Qualitätsaspekte erhält, die man in der jeweiligen Erhebung auch abfragt. Da in der Regel nicht beliebig viele Aspekte erfragt werden können, ist die Auswahl der zu erhebenden Attribute von hoher Bedeutung. Für die Merkmalsauswahl kommen die folgenden Quellen in Betracht:

- „Ableitung aus Unternehmenszielen
- Festlegungen durch das Management
- Qualitative, explorative Befragungen von Kunden (z.B. Kundeninterview, Fahrgastforum, Gruppendiskussion, Panel)

- Expertenbefragungen (sowohl Praktiker aus den Verkehrsunternehmen wie auch Wissenschaftler)
- Literaturrecherchen
- Sichtung der bereits existierenden Instrumente [...]
- Austausch, Kooperation mit anderen Unternehmen
- Analyse eingegangener Beschwerden
- Ergebnis der Arbeit von Qualitätszirkeln, Workshops, Mitarbeiterbefragungen“ (Schnippe 1998, S. 37).

Im Allgemeinen ist es sinnvoll, neben den einzelnen Qualitätseigenschaften ein separates Globalurteil zu erheben. Die Beschränkung auf ein solches Gesamtqualitätsurteil („eindimensionaler Ansatz“) ist zwar vor dem Hintergrund eines adäquaten Qualitätsmanagements nicht empfehlenswert, man kann jedoch davon ausgehen, dass sich in einem solchen Globalurteil zur Zufriedenheit die für den Einzelnen wesentlichen verhaltensrelevanten Aspekte widerspiegeln - auch solche, die über konkrete ÖPNV-bezogene Erfahrungen hinausgehen (z.B. Parkplatzsituation am Arbeitsort und vieles mehr). Darüber hinaus wird einem solchen separaten Globalurteil eine sehr hohe Reliabilität und Validität zugesprochen (vgl. Kroeber-Riel & Weinberg 1996, S. 309). In der vorliegenden Kundenzufriedenheitsbefragung wurden folgende Formulierungen zur Abbildung der Globalzufriedenheit benutzt:

„Wie zufrieden sind Sie ganz allgemein mit dem öffentlichen Nahverkehr in <Gebiet>?“ (einstellungsorientiert).

„Wie zufrieden waren Sie auf dieser Fahrt mit <benutztes Verkehrsmittel> ganz allgemein?“ (zufriedenheitsorientiert).

Bei Nicht-Nutzern:

„Wie beurteilen Sie ganz allgemein den öffentlichen Nahverkehr in <Gebiet>?“

In befragungstechnischer Hinsicht ist es wichtig, die Frage nach dem Globalurteil ganz am Anfang des Fragebogens zu platzieren, damit die Antwort unbeeinflusst von den Einschätzungen der einzelnen Qualitätsmerkmale (Pünktlichkeit, Sauberkeit etc.) ist. Generell sollten Fragen, bei denen sich die Antwort möglicherweise auf die Beantwortung anderer Fragen auswirkt, weiter hinten im Fragebogen platziert werden. Fragt man z.B. Verhalten und Einstellungen gemeinsam ab, so sollten die Fragen zur Einstellung nach denen zum Verhalten gestellt werden (siehe Kapitel 3.3.5).

Antwortskalen

Zur Messung in Form von Rating-Skalen steht eine ganze Reihe von Skalentypen zur Verfügung, die sich grob in numerische (z.B. Zahlen von 1 bis 5), verbale (beschreibende Wörter wie z.B. „sehr“ bis „gar nicht“) und grafische Schätzskalen unterteilen lassen. Beispiel für grafische Skalen sind die Thermometer-Skala, Leiter-Skala oder auch die so genannten Kunin-Items („Smileys“). Darüber hinaus existieren auch Mischformen dieser Typen, wie z.B. numerische Skalen, bei denen nur die beiden Endpunkte mit verbalen Etiketten belegt sind.

Insbesondere bei verbalen Skalen muss bei der Konstruktion entschieden werden, ob unipolare (z.B. „sehr zufrieden“ bis „überhaupt nicht zufrieden“ bzw. „stimme voll und ganz zu“ bis „stimme überhaupt nicht zu“) oder bipolare (z.B. „sehr zufrieden“ bis „sehr unzufrieden“ bzw. „stimme voll und ganz zu“ bis „lehne voll und ganz ab“) Skalen verwendet werden. Bipolare Skalen haben den Vorteil, dass sie für die Befragten oft leichter zu handhaben sind, weil sich die beiden gegenläufigen Begriffe (z.B. sauber - schmutzig) wechselseitig definieren, allerdings ist es manchmal schwierig, zu einem Begriff einen geeigneten Gegenbegriff zu finden.

Eine weitere Frage im Rahmen der Skalenkonstruktion ist die, ob eine gerade oder ungerade Anzahl von Skalenpunkten benutzt werden soll. Bei einer ungeraden Anzahl kann die - meist neutrale - Mittelkategorie zu Verzerrungen führen, da sie bei unsicheren Urteilen als einfache Möglichkeit benutzt wird, sich nicht für die eine oder andere Seite entscheiden zu müssen (obwohl vielleicht durchaus eine Meinung vorhanden ist). Andererseits kann eine gerade Anzahl (was gleichbedeutend mit dem Fehlen einer Mittelkategorie ist) dazu führen, dass Personen, die weder zur einen oder anderen Seite tendieren oder denen der Sachverhalt tatsächlich egal ist, in eine bestimmte Richtung gedrängt werden und so das Gesamtergebnis verfälscht wird. Im Hinblick auf die Zahl der Abstufungen besteht ein Zielkonflikt zwischen dem Wunsch, möglichst fein zu messen und der Gefahr, die Befragten durch eine zu differenzierte Skala zu überfordern. Als Empfehlung kann eine Stufenzahl von 3 bis 7 ausgesprochen werden, wobei hier auch die Befragungsmethodik eine Rolle spielt. Wenn ohne schriftliche Unterlagen gearbeitet wird, wie z.B. in rein telefonischen Interviews, so sollten eher weniger Skalenpunkte verwendet werden, weil der Befragte diese während des Interviews stets im Gedächtnis präsent haben muss. Ansonsten besteht die Gefahr, dass allein aus Gründen der besseren Erinnerung die zuletzt vorgelesenen Antwortkategorien von den Befragten häufiger gewählt werden (sog. „Recency-Effekt“).

In der vorliegenden - rein telefonisch durchgeführten - Kundenzufriedenheitsbefragung wurden 5-stufige Skalen benutzt, bei denen aber nur die beiden Endpunkte mit verbalen Etiketten belegt waren. Es wurde folgende Formulierung gewählt:

„Bitte sagen Sie mir auf einer Skala von 1 bis 5, wie zufrieden Sie damit in <Gebiet> sind. 1 bedeutet sehr zufrieden, 5 bedeutet gar nicht zufrieden. Mit den Werten dazwischen können Sie ihre Meinung abstufen.“

Eine Alternative ist die Verwendung der Schulnotenskala („sehr gut“ bis „ungenügend“), da man davon ausgehen kann, dass praktisch alle Befragten damit vertraut sind und diese somit auch problemlos anwenden können.

Auch die explizite Vorgabe einer „weiß nicht“-Kategorie sollte von der Befragungsmethodik abhängig gemacht werden. In telefonischen und mündlich-persönlichen Interviews empfiehlt es sich, eine „weiß nicht“-Kategorie zwar vorzusehen, diese aber dem Befragten nicht zur Auswahl zu stellen. Andererseits sollte nicht um jeden Preis eine Antwort erzwungen werden, da sonst die Gefahr von Zufallsantworten besteht, was sich negativ auf die Korrelationen mit anderen Merkmalen auswirken kann (zu methodischen Fragen der Skalenkonstruktion vgl. z.B. Trommsdorff 1975 oder Bortz & Döring 2006).

3.3.2 Direkte vs. indirekte Messung

Beim Begriffspaar direkte/ indirekte Messung geht es darum, ob auch die Erwartungen des Kunden abgefragt werden oder nicht. Während bei der direkten Messung nur die wahrgenommene Qualitätseinschätzung (Einstellung, Zufriedenheit oder auch Erwartungsabweichung) bezüglich der einzelnen Merkmale erhoben wird, besteht im Rahmen einer indirekten Messung die Aufgabe des Interviewten darin, einen Ist-Soll-Vergleich bei jedem abgefragten Merkmal vorzunehmen. Als Beispiel für eine solche indirekte Messung kann das Einstellungsmodell von Trommsdorff (1975) genannt werden. Hier wird pro Merkmal die Differenz aus der wahrgenommenen und der aus Sicht des Kunden idealen Ausprägung gebildet. Durch Summation über alle Merkmale erhält man für jeden Befragten die Einstellung gegenüber dem Einstellungsobjekt, das z.B. in einer speziellen Leistung oder auch dem Unternehmen als Ganzes bestehen kann.

Bei der indirekten, zufriedenheitsorientierten Messung kann man nochmals danach unterscheiden, ob die Kundenerwartung vor (ex ante) oder nach (ex post) dem Konsumerlebnis erhoben wird. Ein praktisches Problem bei der ex-ante-Version besteht allerdings darin, dass oftmals im Vorhinein nicht bekannt ist, wer in Zukunft Kunde des Unternehmens sein wird und wer nicht. Insgesamt ist die Erhebung von Kundenerwartungen mit einigen Schwierigkeiten behaftet, da diese auf verschiedenste Art und Weise vom Befragten verstanden werden können (z.B. als ideales-, angemessenes-, aller Erfahrung nach zu erwartendes- oder mindestens zu erwartendes Leistungsniveau; vgl. hierzu z.B. Stauss & Seidel 2006).

3.3.3 Ein- vs. Zweikomponentenansatz

Was schließlich die Unterscheidung in Ein- und Zweikomponentenansatz anbelangt, so findet im Rahmen des Zweikomponentenansatzes zusätzlich eine Gewichtung des jeweiligen Qualitätsmerkmals mit dessen Bedeutung bzw. Wichtigkeit statt. Beim Einkomponentenansatz wird dagegen implizit unterstellt, dass allen Merkmalen die gleiche Bedeutung zukommt. Im Rahmen des Zweikomponentenansatzes können die Bedeutungsgewichte entweder regressionsanalytisch ermittelt oder separat abgefragt werden (zur Vorgehensweise beim ÖPNV-Kundenbarometer vgl. Isfort 2000).

Bei der regressionsanalytischen Variante werden die merkmalsbezogenen Zufriedenheitseinschätzungen als unabhängige Variable und die Globalzufriedenheit als Kriteriumsvariable verwendet. Der Regressionskoeffizient eines jeden Merkmals wird dann als dessen Bedeutungsgewicht interpretiert. Gelegentlich werden dabei die merkmalsbezogenen Zufriedenheiten vorher über eine Faktorenanalyse zu Variablenbündeln zusammengefasst (vgl. z.B. Strasser, Haller & Krietemeyer 2008).

Zur separaten Erhebung von Wichtigkeitseinschätzungen in der Befragung stehen wiederum mehrere Alternativen zur Verfügung. Die gängigste Methode ist dabei die Abbildung über Rating-Skalen (z.B. von „sehr wichtig“ bis „gar nicht wichtig“ bzw. „sehr unwichtig“). Diese Variante wurde auch in der diesem Leitfaden zugrunde liegenden Befragung gewählt:

„Und jetzt möchte ich Sie bitten, mir zu diesen Eigenschaften des öffentlichen Nahverkehrs auf einer Skala von 1 bis 5 zu sagen, wie wichtig Ihnen diese ganz allgemein sind. 1 bedeutet sehr wichtig, 5 bedeutet gar nicht wichtig. Mit den Werten dazwischen können Sie ihre Meinung abstimmen.“

Als Nachteile der Wichtigkeitsermittlung über Rating-Skalen werden häufig interindividuell verschiedene Gewichtungssummen sowie eine Anspruchsinflation („alles ist wichtig“) angeführt (vgl. z.B. Hentschel 1992, S. 145f)¹². Mittels alternativer Techniken wie dem Aufstellen einer Wichtigkeitsrangreihe oder Konstantsummenverfahren (vgl. Schnippe 2000) - also das Verteilen einer festgelegten Zahl von Punkten auf die Merkmale nach Ihrer Wichtigkeit - lassen sich diese Nachteile zwar umgehen, jedoch handelt es sich hier um für die Befragten recht komplexe Auf-

¹² In Tabelle 1 weisen die Durchschnittswerte bei der Wichtigkeit allerdings durchaus eine gewisse Streuung auf. Eine „Anspruchsinflation“ bei Verwendung von Rating-Skalen zur Abfrage der Wichtigkeit von Qualitätseigenschaften („alles ist wichtig“) lässt sich somit auf der Basis der vorliegenden Daten nicht feststellen.

gaben. Diese sind z.B. in telefonischen Interviews aufgrund der fehlenden Visualisierungsmöglichkeiten schwierig anzuwenden, es sei denn, man übermittelt den ausgewählten Befragungspersonen vorab entsprechende Unterlagen.

Bei der Durchführung einer Kundenzufriedenheitsbefragung ist somit im Vorhinein zu entscheiden, ob die Wichtigkeit der interessierenden Qualitätseigenschaften direkt abgefragt oder über eine Regressionsanalyse¹³ aus den Einzelzufriedenheiten und dem Globalurteil berechnet werden. Auf der Basis der Daten der ersten Erhebungswelle wurden beide Vorgehensweisen getestet. In Tabelle 2 wird auf Basis der Teilstichprobe 2 die Rangreihe der ermittelten Regressionskoeffizienten mit der Rangreihe der Mittelwerte bezüglich der abgefragten Wichtigkeiten verglichen:

Tabelle 2: Abgefragte vs. berechnete Wichtigkeit einzelner Merkmale des ÖPNV (alle Gebiete; nur Teilstichprobe 2)

Merkmals	Rang abgefragt	Rang berechnet
Pünktlichkeit	1	5
Sicherheit der Anschlüsse	2	7
Preis-Leistungsverhältnis	3	3
persönliche Sicherheit an Haltestellen	4	10
Anzahl Fahrten/ Verbindungen (Tag)	5	4
Schnelligkeit	6	2
Information an Haltestellen	7	9
Übersichtlichkeit Liniennetz	8	8
Anzahl Fahrten/ Verbindungen (We, Abend)	9	1
Platzangebot im Fahrzeug	10	6
Information im Fahrzeug	11	11

Basis: TSP 2, n=1.290 Befragte

In dem Vergleich der beiden Rangreihen manifestieren sich starke Differenzen zwischen der abgefragten (mittleren) Wichtigkeitsbewertung und der regressionsanalytisch berechneten Wer-

¹³ Regressionsanalyse: Statistisches Analyseverfahren, das den linearen Zusammenhang zwischen einer abhängigen und einer oder mehreren unabhängigen Variablen untersucht. Die betrachteten Variablen müssen metrisches Skalenniveau aufweisen.

tigkeit von Qualitätseigenschaften. Die für die Befragten nach eigener Einschätzung wichtigsten Attribute Pünktlichkeit und Anschlusssicherheit landen bei der regressionsanalytischen Vorgehensweise im Mittelfeld, während die Qualitätseigenschaft mit dem höchsten Regressionskoeffizienten (Anzahl Fahrten/ Verbindungen (Wochenende, Abend)) hinsichtlich des Mittelwertes bei der direkten Abfrage an drittletzter Stelle liegt. Insbesondere wegen der Platzierung der Pünktlichkeit und der Sicherheit der Anschlüsse, die in vielen Studien (vgl. als ein Beispiel die LVB-Kundenzufriedenheitsanalyse 2009/2010; http://www.lvb.de/wir_ueber_uns/Im_Fokus) hinsichtlich der Wichtigkeit weit vorne liegt, erscheint die direkte Abfrage der Wichtigkeit von Qualitätseigenschaften als die plausiblere Variante, auch wenn dies natürlich den Befragungsumfang erhöht.

Bei der Bestimmung der Wichtigkeiten über eine Regressionsanalyse ist zu bedenken, dass dabei alle Einzelzufriedenheiten simultan in die Analyse eingehen und somit der Regressionskoeffizient einer Eigenschaft (als Indikator für die deren Bedeutung) immer auch von der jeweils getroffenen Auswahl an Eigenschaften abhängt. Fragt man die Wichtigkeiten dagegen direkt ab, so ist - zumindest unter Verwendung von Rating-Skalen - die Wichtigkeitseinschätzung der Eigenschaft A weitgehend unabhängig von der Einschätzung der Eigenschaft B.

Auswertung

Insbesondere für investive Entscheidungen ist es von Bedeutung festzustellen, bei welchen Merkmalen Handlungsbedarf besteht. Im Falle eines Zweikomponentenansatzes muss hierfür dann die Zufriedenheitskomponente (bzw. die Soll-Ist-Abweichung bei einer indirekten Messung) und die Wichtigkeitskomponente miteinander verknüpft werden. Wurde die Wichtigkeitseinschätzung über eine Rating-Skala abgefragt, kann z.B. so vorgegangen werden, dass pro Merkmal die beiden Mittelwerte miteinander multipliziert werden. Bei einer entsprechenden Codierung (hoher Wert bei Zufriedenheitseinschätzung = geringe Zufriedenheit; hoher Wert bei Wichtigkeitseinschätzung = große Bedeutung) ergibt sich dann bei dem Merkmal mit dem höchsten Produkt der größte Handlungsbedarf usw.

Bei der Datenanalyse bzw. der Ergebnisdarstellung kann auch ein so genanntes Zufriedenheitsportfolio erstellt werden, in dem jede Qualitätseigenschaft nach (mittlerer) Zufriedenheit und Wichtigkeit (jeweils hoch - niedrig) grafisch in einem zweidimensionalen Diagramm abgetragen wird, wie dies im folgenden Schema angedeutet ist:

Abbildung 6: Portfolio zur Analyse der Kundenzufriedenheit

Als Kriterium zur Trennung der beiden Komponenten kann z.B. jeweils der Median der Durchschnittswerte herangezogen werden (für andere Varianten vgl. Kriemeyer & Oelbaum 1997). Selbstverständlich sind auch feinere Differenzierungen des Zufriedenheits-Portfolios möglich (vgl. z.B. Schnippe 1998). Benutzt man zur Trennung von hoher bzw. niedriger Zufriedenheit (Wichtigkeit) nicht externe Vergleichsgrößen, sondern die Werte aus derselben Erhebung, so muss man sich darüber im Klaren sein, dass hieraus immer nur ein relativer, aber kein „absoluter“ Handlungsbedarf abgeleitet werden kann.

Eine weitere Analysemöglichkeit ist die Anwendung einer Penalty-Reward-Faktoren-Analyse, bei der es im Hinblick auf Maßnahmenentwicklung insbesondere um das Aufdecken von Penalty-Faktoren geht. Hierauf wird im nächsten Abschnitt eingegangen.

3.3.4 Penalty-Reward-Faktoren-Analyse

Bei der Penalty-Reward-Faktoren-Analyse (PRFA) handelt es sich um ein eigenständiges merkmalsorientiertes Verfahren (neben der multiattributiven Messung und dem Conjoint Measurement). Da jedoch die Daten aus einer multiattributiven Messung in aller Regel (es muss eine Globalzufriedenheit vorhanden sein) mittels einer PRFA ausgewertet werden können, wird dieses Verfahren hier im Kontext der multiattributiven Messung behandelt.

Bei der PRFA geht es um Zusammenhänge zwischen Globalzufriedenheit und der Zufriedenheit mit einzelnen Qualitätsmerkmalen. Anders als in der „normalen“ Regressionsanalyse wird hier

jedoch kein linearer Ansatz verwendet. Vielmehr werden dabei die Einzelzufriedenheiten in so genannte „dummy-Variable“ (0/1-Merkmale) zerlegt, mit denen sich nicht-lineare Zusammenhänge aufzeigen lassen. Wurden die Qualitätseigenschaften z.B. auf einer 5-stufigen Skala gemessen, so kann jedes Zufriedenheitsmerkmal (z.B. Pünktlichkeit) nach dem folgenden Schema in zwei dummy-Variable zerlegt werden:

„Pünktlichkeit dummyA“ = 1, falls Antwort bei Pünktlichkeit = 1 oder 2 (hohe Zufriedenheit)

„Pünktlichkeit dummyA“ = 0 sonst

„Pünktlichkeit dummyB“ = 1, falls Antwort bei Pünktlichkeit = 4 oder 5 (niedrige Zufriedenheit)

„Pünktlichkeit dummyB“ = 0 sonst

Für die Ausprägung 3 von Pünktlichkeit darf keine dummy-Variable gebildet werden, diese stellt die so genannte Referenzkategorie dar. Die erhobenen und solchermaßen in 0/1-Merkmale zerlegten Qualitätseigenschaften gehen als unabhängige Variable in ein Varianzanalysemodell ein, in dem die Globalzufriedenheit als abhängige Größe fungiert. In der folgenden Tabelle sind die aus der Teilstichprobe 2 der durchgeführten Kundenzufriedenheitsbefragung resultierenden Ergebnisse einer solchen Analyse (alle Gebiete) zusammengestellt:

Tabelle 3: Ergebnisse der Penalty-Reward-Faktoren-Analyse (alle Gebiete; nur Teilstichprobe 2)

Merkmals	dummy-Variable	Schätzwert	Signifikanz
Anzahl Fahrten/ Verbindungen (Tag)	niedrige Zufr.	0,21	0,0933
	hohe Zufr.	-0,26	0,0004
	Referenz	0	-
Anzahl Fahrten/ Verbindungen (We, Abend)	niedrige Zufr.	0,31	<0,0001
	hohe Zufr.	-0,19	0,0041
	Referenz	0	-
Sicherheit der Anschlüsse	niedrige Zufr.	0,02	0,8265
	hohe Zufr.	-0,12	0,0679
	Referenz	0	-
Pünktlichkeit	niedrige Zufr.	0,18	0,0722
	hohe Zufr.	-0,16	0,0112
	Referenz	0	-
Schnelligkeit	niedrige Zufr.	0,16	0,1659
	hohe Zufr.	-0,19	0,0031
	Referenz	0	-
Übersichtlichkeit Liniennetz	niedrige Zufr.	0,30	0,0018
	hohe Zufr.	-0,11	0,0834
	Referenz	0	-
Platzangebot im Fahrzeug	niedrige Zufr.	0,37	<0,0001
	hohe Zufr.	-0,03	0,6294
	Referenz	0	-
Information an Haltestellen	niedrige Zufr.	0,11	0,1824
	hohe Zufr.	-0,06	0,3806
	Referenz	0	-
Information im Fahrzeug	niedrige Zufr.	0	0,9662
	hohe Zufr.	-0,02	0,7400
	Referenz	0	-
Preis-Leistungsverhältnis	niedrige Zufr.	0,17	0,0064
	hohe Zufr.	-0,22	0,0011
	Referenz	0	-
persönliche Sicherheit an Haltestellen	niedrige Zufr.	0,17	0,0295
	hohe Zufr.	0,07	0,2516
	Referenz	0	-

Basis: TSP 2, n=1.290 Befragte

Nach diesen Ergebnissen wären die Merkmale Sicherheit an Haltestellen, Übersichtlichkeit des Liniennetzes und insbesondere Platzangebot im Fahrzeug als Penalty-Faktoren zu bewerten. Hier führt - im Vergleich zur Referenzkategorie (Ausprägung 3) - eine niedrige Zufriedenheit zu einem signifikant höheren Mittelwert bei der Globalzufriedenheit (was bei der gewählten Codierung eine niedrigere Globalzufriedenheit bedeutet), während eine hohe Zufriedenheit sich **nicht** (signifikant) auf die Globalzufriedenheit auswirkt. Betrachtet man dagegen das Merkmal Preis-Leistungsverhältnis, so führt eine niedrige Zufriedenheit zu einem signifikant höheren Mittelwert der abhängigen Variable und umgekehrt; hier handelt es sich somit weder um einen Penalty- noch um einen Reward-Faktor. Merkmale mit einer eindeutigen Reward-Struktur lassen sich im

vorliegenden Beispiel nicht identifizieren. Wie auch die vorliegenden Auswertungen zeigen, sind die aus einer PRFA resultierenden Ergebnisse nicht immer eindeutig interpretierbar.

3.3.5 Messung der Kundenbindung

Neben der Zufriedenheit der Kunden mit der Angebotsqualität ist auch die Bindung der Kunden an das Unternehmen eine wichtige Größe im Marketing. Wenngleich die Kundenbindung nicht in direktem Zusammenhang mit der multiattributiven Zufriedenheitsmessung steht, soll an dieser Stelle kurz auf das Thema Kundenbindung eingegangen werden (zur Messung der Kundenbindung vgl. z.B. Homburg & Fürst 2008).

In der Befragung wurde die Kundenbindung über die folgenden 4 Fragen abgebildet (nur ÖPNV-Nutzer):

- *„Wie häufig werden Sie den öffentlichen Nahverkehr in <Gebiet> zukünftig nutzen?“*
(genauso oft wie heute; häufiger als heute; seltener als heute; gar nicht mehr)
- *„Wie oft haben Sie den öffentlichen Nahverkehr in <Gebiet> in der Vergangenheit an Freunde, Bekannte oder Verwandte weiterempfohlen?“*
(noch nie; selten; gelegentlich; schon häufig)
- *„Beabsichtigen Sie, den öffentlichen Nahverkehr in <Gebiet> zukünftig an Freunde, Bekannte oder Verwandte weiterzuempfehlen?“*
(ja; nein)
- *„Zu Beginn unseres Gesprächs sagten Sie, dass Sie den öffentlichen Nahverkehr zu den Zwecken <...> nutzen. Könnten Sie sich darüber hinaus vorstellen, den öffentlichen Nahverkehr zukünftig noch für andere Zwecke zu nutzen?“*
(ja; nein)

Die erste Frage bildet die zukünftige Nutzungs-/ (Wiederkauf-)absicht, die zweite das Weiterempfehlungsverhalten und die dritte die Weiterempfehlungsabsicht ab. In der vierten Frage spiegelt sich die Zusatzkaufabsicht wider, wobei hier auf eine vorher gestellte Frage nach den aktuell genutzten Fahrtzwecken Bezug genommen wird.

Diese Indikatoren der Kundenbindung können für sich alleine betrachtet oder zu einem Indexwert zusammengefasst werden. Eine mögliche Variante für die Bildung eines solchen Kundenbindungsindex' ist in der folgenden Abbildung beispielhaft dargestellt:

Abbildung 7: Konstruktion eines Kundenbindungsindex'

Variable	Ausprägungen der Variable	Punktzahl	
Zukünftige Nutzungshäufigkeit des ÖPNV	1 genauso oft wie heute	1	1
	2 häufiger als heute	1	
	3 seltener als heute	0	
	4 gar nicht mehr	0	
Weiterempfehlung in der Vergangenheit	1 noch nie	0	0
	2 selten	0	
	3 gelegentlich	1	
	4 schon häufig	1	
Weiterempfehlung in der Zukunft	1 ja	1	1
	2 nein	0	
Zukünftige Nutzung für andere Zwecke	1 ja	1	0
	2 nein	0	

Kundenbindungsindex: (keine) 0 1 2 3 4 (maximale) **2**

←—————→

In diesem Beispiel werden pro Merkmal die Ausprägungen, die auf eine hohe Kundenbindung hindeuten, mit 1, die anderen mit 0 codiert. Der Kundenbindungsindex wird dann aus der Summe dieser 0/1-Merkmale gebildet (additiver Index), d.h., die resultierende Skala liegt im Bereich zwischen 0 und 4. In dem in Abbildung 7 dargestellten Beispiel ergibt sich ein Indexwert von 2.

3.4 Critical Incident-Verfahren

3.4.1 Erhebung

Wie bereits erwähnt, hebt die Critical Incident-Technik auf besondere Vorfälle im Dienstleistungsprozess ab, die vom Kunden als außergewöhnlich positiv oder negativ empfunden wurden.

Zur Erhebung der kritischen Ereignisse können die folgenden Fragen dienen:

- „Hatten Sie persönlich schon einmal ein besonders gutes Erlebnis mit dem öffentlichen Nahverkehr in <Gebiet>?“ (ja/ nein); falls ja:
 - „Bitte schildern Sie es mir kurz“ (offene Frage)
 - „Haben Sie Freunden, Verwandten oder Bekannten von diesem positiven Erlebnis

erzählt?“ (ja/ nein)

falls ja: „Wie vielen Personen haben Sie ungefähr davon erzählt?“

- *„Hatten Sie persönlich schon einmal ein besonders schlechtes Erlebnis mit dem öffentlichen Nahverkehr in <Gebiet>?“ (ja/ nein); falls ja:
 - *„Bitte schildern Sie es mir kurz“ (offene Frage)*
 - *„Haben Sie Freunden, Verwandten oder Bekannten von diesem negativen Erlebnis erzählt?“ (ja/ nein)*
 - falls ja: „Wie vielen Personen haben Sie ungefähr davon erzählt?“*
 - *„Haben Sie sich über dieses schlechte Erlebnis beschwert?“ (ja/ nein)**

Neben der eigentlichen Erhebung der Ereignisse, für die keine Antwortvorgaben gemacht wurden (offene Fragen), wurde somit zusätzlich noch nach dem jeweiligen Weitererzählungsverhalten und - bei negativen Ereignissen - nach dem Beschwerdeverhalten gefragt.

3.4.2 Auswertung

Als Ergebnis dieser offenen Fragen erhält man eine breite Palette unterschiedlichster Aussagen und Statements. Um diese quantitativ beschreibbar zu machen, muss unter Verwendung inhaltsanalytischer Methoden ein Kategorienschema erstellt werden, das eine überschneidungsfreie und vollständige Zuordnung aller Nennungen erlaubt. Auf der Basis des aus der Kundenzufriedenheitsbefragung vorliegenden Datenmaterials wurden insgesamt 15 Grobkategorien unterschieden, die jeweils nochmals in Unterkategorien ausdifferenziert sind. In der folgenden Abbildung ist zur Illustration ein Auszug aus dem Kategorienschema dargestellt:

Abbildung 8: Kategorienschema zur Klassifizierung der kritischen Ereignisse (Auszug)

Codeplan für offene Fragen

Code		Kategorien		Erwartete Besetzung		
grob	fein	grob	fein	Positives Erlebnis	Negatives Erlebnis	Veränderung
10	10	Andere Fahrgäste	Andere Fahrgäste: allgemein	X	X	X
	11		Andere Fahrgäste: Begegnungen, soziale Kontakte	X		
	12		Andere Fahrgäste: Hilfsbereitschaft	X	X	X
	13		Andere Fahrgäste: Störung durch asoziales Verhalten		X	X
	14		Andere Fahrgäste: Ausländerfeindlichkeit		X	X
	15	Andere Fahrgäste: pöbelnde/lärmende Jugendliche		X	X	
20	20	Anzahl Fahrten	Anzahl Fahrten: allgemein	X	X	X
	21		Anzahl Fahrten: Betriebszeit (abends, We), z.B. Nachtbus	X	X	X
	22		Anzahl Fahrten: Betriebszeit (tagsüber)	X	X	X
	23		Anzahl Fahrten: Takt (abends, We)	X	X	X
	24	Anzahl Fahrten: Takt (tagsüber)		X	X	
30	30	Fahrzeug	Fahrzeug: allgemein	X	X	X
	31		Fahrzeug: neu, modern, komfortabel (klimatisierung)	X	X	X
	32		Fahrzeug: Platzangebot/Besetzungsgrad	X	X	X
	33		Fahrzeug: Werbung		X	X
	34		Fahrzeug: Umweltfreundlichkeit	X		X
	35		Fahrzeug: technische Probleme/Störung		X	X
40	40	Fahrgastinformation	Information: allgemein	X	X	X
	41		Information: an Haltestellen	X	X	X
	42		Information: im Fahrzeug	X	X	X
	43		Information: im Störfall	X	X	X

Anmerkung: Die Spalte „Veränderung“ bezieht sich auf die offene Frage „Hat sich nach Ihrer Wahrnehmung in der letzten Zeit rund um den öffentlichen Nahverkehr in <Gebiet> irgendetwas verändert?“. Auch hier mussten die Nennungen entsprechend kategorisiert werden.

3.4.3 Ergebnisse

Üblicherweise wird die Critical Incident-Technik eher im Rahmen von qualitativen (in der Regel mündlich-persönlichen) Interviews benutzt, in dem die befragten Personen in Form von offenen Fragen breiten Raum haben, ihre Erlebnisse mit den entsprechenden Begleitumständen sowie die damit verbundenen Erfahrungen und Empfindungen möglichst detailliert zu beschreiben. Oft werden diese Interviews auch per Tonband aufgezeichnet. Im Rahmen des vorliegenden Forschungsvorhabens wurde der Versuch unternommen, die Critical Incident-Technik im Kontext einer telefonischen, standardisierten Befragung anzuwenden, was sich im Großen und Ganzen als praktikabel erwiesen hat. Allerdings müssen die Interviewer die Antworten auf die offenen Fragen während des Interviews erfassen, was unter Umständen zu Verkürzungen und - was soweit wie möglich vermieden werden sollte - zu Interpretationen von Seiten der Interviewer führen kann. Als wichtig hat sich darüber hinaus herausgestellt, die Befragten dazu anzuhalten, die einzelnen Ereignisse sauber voneinander zu trennen. Insgesamt erfordert die Anwendung der Critical Incident-Technik in einer telefonischen Befragung eine gute Schulung der Interviewer sowie einen relativ hohen Aufwand für die Bildung eines Kategorienschemas und die Einordnung der einzelnen Angaben in dieses Schema.

Die aus den Daten der ersten Erhebungswelle (mit n=3.153 Befragten) resultierenden Ergebnisse zeigen, dass sich die geschilderten kritischen Ereignisse sehr häufig auf Kontakte bzw. Konflikte mit anderen Personen (Personal oder andere Fahrgäste) beziehen. Bei den negativen Erlebnissen entfallen rund 45 % der Nennungen auf diese beiden Bereiche:

Tabelle 4: Verteilung der positiven und negativen Erlebnisse (grobe Kategorisierung)

Positive Erlebnisse	Antworten	Anteil	Negative Erlebnisse	Antworten	Anteil
Personal	400	61,3 %	Personal	471	34,9 %
Andere Fahrgäste	59	9,0 %	Pünktlichkeit	248	18,4 %
Service	39	5,9 %	Andere Fahrgäste	140	10,4 %
Pünktlichkeit	37	5,6 %	persönliche Sicherheit	106	7,9 %
Anzahl der Fahrten	18	2,8 %	Fahrzeug	97	7,2 %
Sicherheit der Anschlüsse	18	2,7 %	Information	85	6,3 %
Information	12	1,8 %	Infrastruktur	48	3,5 %
Schnelligkeit der Beförderung	11	1,6 %	Sicherheit der Anschlüsse	39	2,9 %
Fahrzeug	10	1,6 %	Service	34	2,5 %
Linienetz	9	1,4 %	Anzahl der Fahrten	21	1,6 %
Tarif	7	1,1 %	Linienetz	15	1,1 %
Infrastruktur	6	0,9 %	Sauberkeit	12	0,9 %
persönliche Sicherheit	6	0,9 %	Schnelligkeit der Beförderung	9	0,7 %
Sauberkeit	2	0,2 %	Tarif	7	0,5 %
Sonstige	14	2,2 %	Umsteigezeit (im Verhältnis zur Fahrtzeit)	5	0,3 %
Zuordnung nicht möglich	6	0,9 %	Sonstige	6	0,5 %
			Zuordnung nicht möglich	5	0,4 %
Gesamt*	652	100 %	Gesamt*	1349	100 %

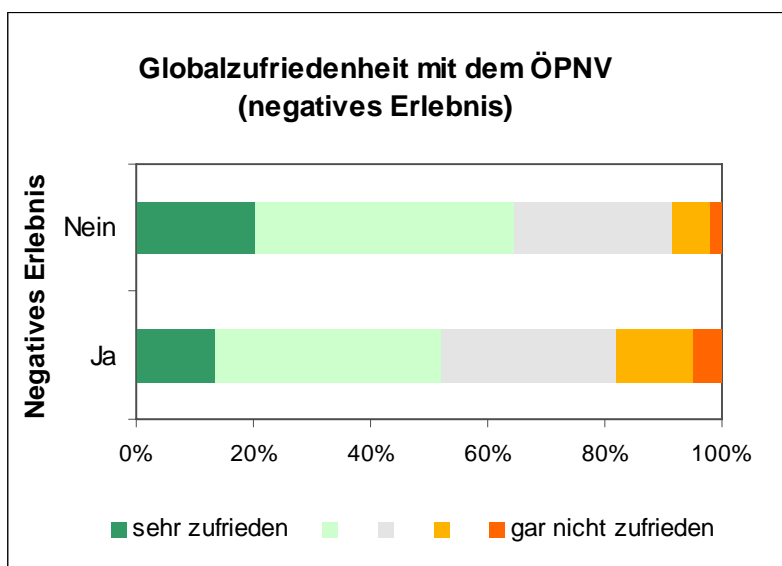
*Abweichungen der Gesamtwerte von der Summe der Einzelwerte sind auf Rundungen zurückzuführen.
Basis: 630 Befragte (MFN) Basis: 1.118 Befragte (MFN)

Andere Qualitätsaspekte im Zusammenhang mit der ÖPNV-Nutzung, die für die Kunden jedoch durchaus problematisch - und ggf. verhaltensrelevant - sein können (wie z.B. Fahrtenhäufigkeit oder Preise), werden dagegen eher selten berichtet, weil diese im Kontext der offenen Frageformulierung offensichtlich nicht als besondere, singuläre Ereignisse wahrgenommen werden.

Die Frage ist, ob sich allein aus solchen Critical Incidents valide Rückschlüsse auf die Kundenzufriedenheit ziehen lassen, zumal das ein oder andere geschilderte Erlebnis unter Umständen schon länger zurück liegen kann (wie z.B. das Statement „Abfahrt einer Straßenbahn ohne Aufnahme von Passagieren 1976“!). Dabei kann es durchaus sein, dass in Einzelfällen z.B. ein negatives Erlebnis dazu geführt hat, dass der ÖPNV nicht mehr benutzt wird. Deshalb stellt sich natürlich die Frage nach der Verhaltensrelevanz und den Auswirkungen solcher Critical Incidents auf die Kundenzufriedenheit. Denkbar wäre schließlich auch, dass sich gute und schlechte Erfahrungen bezüglich ihrer Wirkung auf das Verhalten (Art und Häufigkeit der ÖPNV-Nutzung) ggf. gegenseitig neutralisieren.

Zum Abschluss soll deshalb der Zusammenhang zwischen Kundenzufriedenheit und Critical Incidents beleuchtet werden. Hierfür wird die Globalzufriedenheit von Personen mit einem negativen Erlebnis mit der von Personen ohne ein solches Erlebnis verglichen.

Abbildung 9: Kundenzufriedenheit und negative Erlebnisse



Basis: n=3.076

Wie in Abbildung 9 erkennbar ist, sind die Unterschiede zwischen den beiden Gruppen recht deutlich: Personen, die (mindestens) ein negatives Erlebnis berichtet haben, sind unzufriedener als Personen die noch kein negatives Erlebnis mit dem ÖPNV hatten. Da sich über 70 % der Nennungen bei den negativen Erlebnissen auf die Bereiche Personal, Pünktlichkeit, andere Fahrgäste und persönliche Sicherheit beziehen (siehe Tabelle 4), ist davon auszugehen, dass insbesondere von Verbesserungen hinsichtlich der Ausbildung des Personals (hier vor allem des Fahrpersonals), der Pünktlichkeit und des Einsatzes von Sicherheitspersonal (wegen „andere Fahrgäste“ und „persönliche Sicherheit“) eine Steigerung der Kundenzufriedenheit erwartet werden kann.

4 Bewertung der Verfahren hinsichtlich der Möglichkeiten zur Ableitung von Maßnahmen

Folgt man dem Ansatz, dass sich Maßnahmen im öffentlichen Verkehr an der Zufriedenheit von (potenziellen) Kunden mit der Leistungsqualität orientieren sollten und akzeptiert man die Definition von Kundenzufriedenheit als Vergleich von erwarteter und wahrgenommener Qualität, so kommt man zu dem Schluss, dass die Messverfahren, die auf einer - möglichst repräsentativen - Befragung von Personen im Bedienungsgebiet basieren, am ehesten zur Ableitung entsprechender Maßnahmen geeignet sind.

Verfahren wie *direkte Leistungsmessung* oder *Beobachtungs- bzw. Testkundenverfahren* haben zwar durchaus ihren Stellenwert, wenn es z.B. darum geht, im Rahmen von Bonus-/ Malus-Vereinbarungen die vom Anbieter angestrebte Qualität mit dem tatsächlich erbrachten Leistungsniveau abzugleichen. Hier bleibt jedoch die Frage, ob sich die vom Anbieter angestrebte Qualität mit der von den Kunden erwarteten Leistung deckt (eine gewisse Ausnahme bilden hier Testkundenverfahren, bei denen „echte Kunden“ als Tester eingesetzt werden). Sind die Kundenerwartungen höher, führt auch die Einhaltung der vom Unternehmen festgelegten Qualitätsziele nicht zur Zufriedenheit bei den Kunden. Auf der anderen Seite können bei Nichterfüllung der Qualitätsziele unter Umständen Maßnahmen initiiert werden, die aus Kundensicht gar nicht erforderlich sind, weil in dem Bereich, auf den sich die Maßnahmen beziehen, bereits eine hohe Zufriedenheit besteht.

Auch in der *Beschwerdeanalyse* kann ein hilfreiches Instrument zur Entdeckung von Schwachstellen im Prozess der Dienstleistungserbringung gesehen werden. Die Problematik dabei besteht darin, dass sich in der Regel nur ein Bruchteil der unzufriedenen Kunden beim Unternehmen beschwert und die Personen, die eine Beschwerde vorbringen, mit hoher Wahrscheinlichkeit nicht repräsentativ für die Gesamtheit der Kunden mit negativen Erlebnissen sind. Die Frage nach der Repräsentativität der involvierten Kunden stellt sich auch bei *Kundengremien* oder *-foren*, so dass eine Ableitung von Maßnahmen allein auf der Basis dieser Verfahren unter Umständen zu einer Fehlallokation von Mitteln führen kann.

Aus den genannten Gründen erscheint somit die Befragung von (potenziellen) Kunden als die am besten geeignete Erhebungsform für die Maßnahmenentwicklung. Hierbei stellt sich dann die Frage nach einer merkmals- oder eher ereignisorientierten Vorgehensweise.

Die Stärke der *ereignisorientierten Verfahren* liegt darin, dass man damit sehr detaillierte Informationen zum Prozess der Dienstleistungserstellung oder zu kritischen Ereignissen erzielen kann. Ist der Kunde mit einem bestimmten Sachverhalt (z.B. Wartezeit am Verkaufsschalter)

unzufrieden, lassen sich mit diesen Methoden Erkenntnisse über die konkreten Umstände und Ursachen dieser Unzufriedenheit (z.B. ständiges gestört werden der Mitarbeiter durch Telefonanrufe) gewinnen. In der vertieften Analyse solcher „auffälligen“ Sachverhalte sowie auch in der Entwicklung der Erhebungsinstrumente für eine multiattributive Messung kann der spezifische Nutzen der ereignisorientierten Verfahren gesehen werden. Diese sind somit gut zur Entdeckung von Schwachstellen im Angebot und zur Ableitung von Maßnahmen geeignet, wobei allerdings sowohl in der eigenen Befragung sowie auch bei anderen Anwendungen der Critical Incident-Technik festgestellt wurde, dass „normale“ Unzufriedenheit erzeugende Bedingungen wie z.B. Tarif kaum genannt werden (vgl. Probst, Kahrs & Rieschick-Dziambas 2001). Der spezifische Nutzen der Critical Incident-Technik kann somit vor allem darin gesehen werden, Stärken und Schwächen im Zusammenhang mit Kundenkontakten der Mitarbeiter zu identifizieren. Offensichtlich sind es gerade die Erfahrungen bei Kontakten mit Mitarbeitern, welche den Kunden bei einer entsprechenden offenen Frage nach besonderen Erlebnissen im ÖPNV in den Sinn kommen. Auf dieser Basis lassen sich Ansätze zur Entwicklung entsprechender Maßnahmen vor allem im Bereich Kundenumgang ableiten, die dann - da negative Ereignisse häufiger weitererzählt werden als positive - aufgrund der Mund-zu-Mund-Propaganda auch zur Verbesserung des Images beitragen können. Allerdings erfordern diese Verfahren aufgrund ihrer grundsätzlich qualitativen Ausrichtung (offene Fragen; unstrukturierte Interviews) einen hohen Erhebungs- und Analyseaufwand, weshalb oft nur kleine Fallzahlen realisiert werden können. Letztlich beziehen sich diese Methoden auf bestehende Kunden mit konkreten Nutzungserfahrungen, während sie in Bezug auf potenzielle Kunden (ÖV-Nicht-Nutzer) weniger relevant sind.

Bei den *merkmalsorientierten Befragungsmethoden* sind die *Conjoint-Ansätze* zwar nicht unmittelbar auf das Konzept der Kundenzufriedenheit zugeschnitten, sie haben aber doch ihren Stellenwert bei der Maßnahmengestaltung. So lässt sich mit diesen Verfahren z.B. der hypothetische Nutzen eines bislang noch nicht am Markt eingeführten Angebots ermitteln oder es kann eine Entscheidung zwischen verschiedenen Angebotsoptionen getroffen werden. Ein Beispiel hierfür wäre die Fragestellung, ob eine Fahrzeitverkürzung um 5 Minuten einen höheren bzw. geringeren Nutzen für den Kunden stiftet als eine Taktverdichtung von 15 auf 10 Minuten (vgl. Bogner & Mayer 1994). Der Nachteil dieser Verfahren besteht darin, dass aufgrund der Komplexität des Erhebungsvorgangs meist nur wenige Merkmale simultan betrachtet werden können.

Insbesondere wenn eine Zeitreihe der Kundenzufriedenheit aufgebaut werden soll oder eine Wirkungskontrolle von Maßnahmen angestrebt wird, ist die Anwendung *multiattributiver Verfahren* aufgrund ihres hohen Grades an Standardisierung und der relativ einfachen Handhabbarkeit sinnvoll. Diese können sowohl in Bezug auf konkrete Fahrten (z.B. Befragung im Fahrzeug) als auch allgemein (und somit auch bei Nicht-Kunden) angewendet werden. Bei Befragungen

im Fahrzeug ist auf eine repräsentative Auswahl von Wagenfahrten zu achten, da bei einer Beschränkung auf einige wenige Fahrten möglicherweise Verzerrungen auftreten, die zu falschen Schlussfolgerungen hinsichtlich der zu ergreifenden Maßnahmen führen können. Der hohe Grad an Standardisierung hat jedoch einerseits zur Konsequenz, dass die abzufragenden relevanten Merkmale (Qualitätseigenschaften) im Vorhinein bekannt sein bzw. festgelegt werden müssen. Zum anderen erhält man im Vergleich zu ereignisorientierten Verfahren relativ pauschale Qualitätseinschätzungen.

Dennoch lassen sich auch mit multiattributiven Verfahren Hinweise auf Schwachstellen im Leistungsangebot und damit auf notwendige Maßnahmen erzielen, z.B. mit Hilfe eines Zufriedenheits-Portfolios. Eine andere Möglichkeit ist die Durchführung einer *Penalty-Reward-Faktoren-Analyse*, die sich in aller Regel mit den im Rahmen einer multiattributiven Messung gewonnenen Daten rechnen lässt. Gelingt es hier, vor allem die Penalty-Faktoren (auch Basisfaktoren genannt) zu identifizieren - also die Mindestanforderungen, die im Fall der Nicht-Erfüllung bei den Kunden Unzufriedenheit hervorrufen -, lassen sich auf dieser Basis Ansatzpunkte für gezielte Strategien ableiten.

Gut geeignet für die Entwicklung von Maßnahmen - aber auch mit einem hohen Aufwand verbunden - ist die Kombination von merkmalsorientierten, insbesondere multiattributiven Methoden einerseits und ereignisorientierten Verfahren - evtl. in verkürzter Form - andererseits. Ein Beispiel für einen solchen Methodenmix lässt sich der Arbeit von Brög und Kahn (2003) entnehmen. Die dort geschilderte Vorgehensweise besteht in der Anwendung einer multiattributiven Messung ohne Bezug auf eine konkrete Fahrt (dort „Imagemessung“ genannt). Diese wird ergänzt durch eine Art sequentielle Ereignisanalyse mit Bezug auf vorher erhobene ÖPNV-Fahrten, wobei im Anschluss versucht wird, daraus die Merkmale der multiattributiven Messung nachzubilden, um Vergleiche zu ermöglichen. Ist nun die Zufriedenheit z.B. mit dem Fahrzeugkomfort bei der Imagemessung höher als bei der ereignisorientierten Messung, spricht dies für betriebliche Maßnahmen. Im umgekehrten Fall sind nach Ansicht der Autoren kommunikative Maßnahmen angeraten.

Eine andere Möglichkeit der Verknüpfung könnte darin bestehen, zuerst im Rahmen eines multiattributiven Ansatzes eine Reihe von Qualitätseinschätzungen abzufragen und anschließend bei schlecht bewerteten Eigenschaften offen nachzufragen, ob es kritische Ereignisse oder allgemein Gründe gibt, die zu dieser Abwertung geführt haben. Auf diese Weise können ggf. konkrete Hinweise auf Verbesserungspotenziale gefunden werden, die durch die reine Anwendung von multiattributiven Verfahren nicht erkannt werden könnten. Voraussetzung hierfür ist die Festlegung von Kriterien, ab wann eine Bewertung als „schlecht“ einzustufen ist.

Generell besteht bei der Maßnahmenableitung das Problem, ab welchem Ausmaß an (Un-)Zufriedenheit überhaupt Maßnahmen zu ergreifen sind (Schwellenwertproblematik). Bei multiattributiven Messungen wird im Rahmen von Zufriedenheits-Portfolios („Handlungsrelevanz-Matrix“) meist der Median oder Mittelwert der Durchschnittswerte herangezogen, um die einzelnen Qualitätseinschätzungen nach hoher/ niedriger Zufriedenheit bzw. hoher/ niedriger Wichtigkeit zu trennen (siehe Abbildung 6). Hierbei handelt es sich jedoch um ein relatives Maß. Weisen z.B. alle Qualitätsmerkmale eine relativ hohe mittlere Zufriedenheit auf, wird nur zwischen mehr oder weniger gut, im umgekehrten Fall nur zwischen mehr oder weniger schlecht differenziert. In beiden Fällen würde man aber auf dieser Basis Handlungsbedarf ableiten, und zwar bei den Merkmalen, die jeweils die geringste durchschnittliche Zufriedenheit aufweisen. Eine bessere Variante, die Merkmale nach hoher/ niedriger Zufriedenheit (und ggf. Wichtigkeit) zu trennen, ist die Verwendung externer Vergleichsinformationen (z.B. von ähnlich strukturierten Unternehmen oder Branchendurchschnitt; vgl. Krietemeyer & Oelbaum 1997, S. 24). Hier entsteht dann allerdings meist das Problem mangelnder Vergleichbarkeit sowohl hinsichtlich der Struktur der zugrunde liegenden Gebiete als auch wegen der Verwendung unterschiedlicher Skalen, Merkmale und/oder Befragungsformen.

In den folgenden beiden Übersichten sind die einzelnen Methoden der Qualitätsmessung nochmals mit ihren Anwendungsgebieten und der Eignung zur Maßnahmenableitung dargestellt.

Tabelle 5: Übersicht über die Methoden der Qualitätsmessung

Methode	Stichprobe	Aufwand	Schwächen	Anwendungsgebiet	Eignung zur Maßnahmenableitung
Direkte Leistungsmessung	-	bei ausschließlicher Verwendung prozessproduzierter Daten eher geringer Aufwand	Schluss auf Kundenzufriedenheit fraglich	Bonus-/ Malus-Regelungen	nur in wenigen Teilbereichen (z.B. Pünktlichkeit)
Beobachtung	Auswahl von Kundenkontakten (räumlich-zeitlich)	relativ hoher Analyseaufwand (Gedächtnisprotokolle); Ausnahme: stark strukturierte Beobachtung (vorgefertigtes Kategorienschema)	Schluss auf Kundenzufriedenheit fraglich	Bonus-/ Malus-Regelungen	nur in Teilbereichen (z.B. in Bezug auf persönliche Kundenkontakte)
Testkundenverfahren	Auswahl von Haltestellen, Automaten, Fahrten etc. (räumlich-zeitlich)	relativ hoher Aufwand für repräsentative Stichprobenauswahl, Schulung der Tester	Schluss auf Kundenzufriedenheit fraglich	Bonus-/ Malus-Regelungen	nur auf Basis objektiver Kriterien
Beschwerdeanalyse	-	relativ hoher Analyseaufwand für systematische Erfassung und Strukturierung der eingehenden Beschwerden	keine Repräsentativität bzgl. der Beschwerdeführer und Art der Beschwerde	Schwachstellenanalyse/ Verbesserungsvorschläge	mit Einschränkungen
Kundengremien	Gefahr von Selbstselektionseffekten bei öffentlichen Aufrufen zur Teilnahme	vergleichsweise gering	Sind die Gremienmitglieder repräsentativ für die große Masse der Kunden?	Schwachstellenanalyse/ Verbesserungsvorschläge	mit Einschränkungen

Tabelle 6: Übersicht über die Methoden der Kundenzufriedenheitsmessung

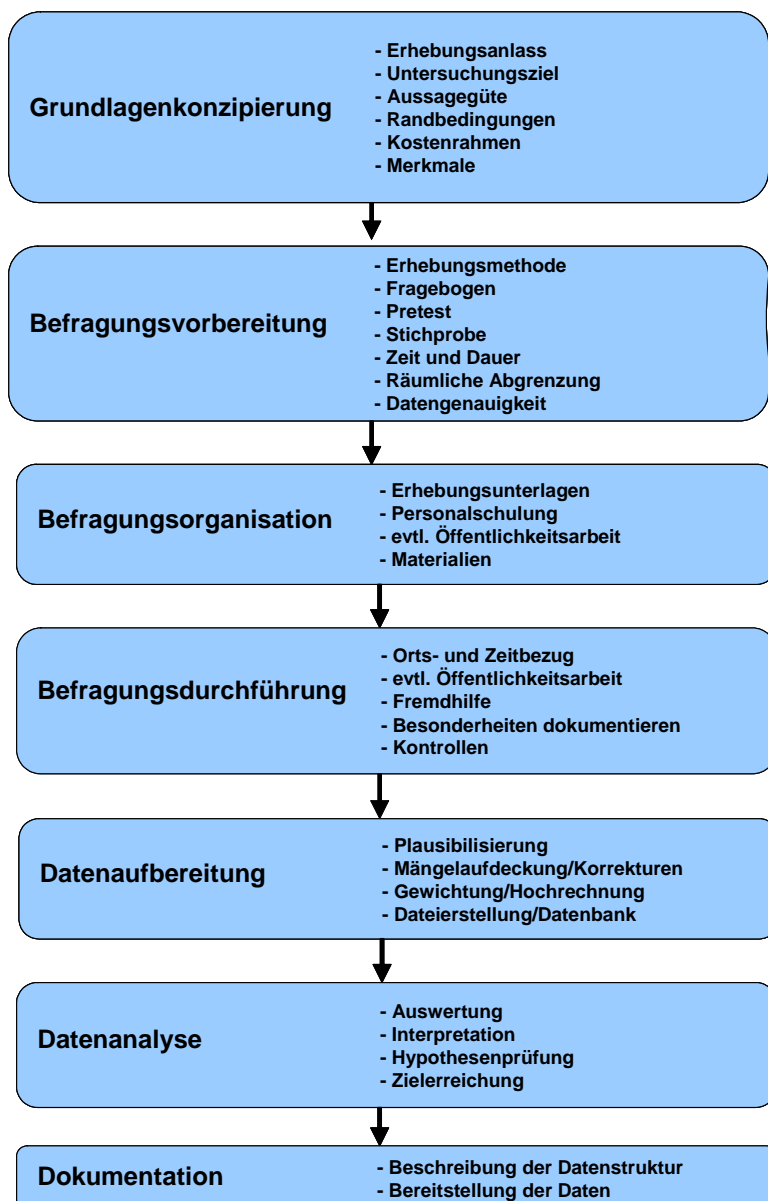
Methoden	Stichprobe	Aufwand	Schwächen	Anwendungsgebiet	Eignung zur Maßnahmenableitung
Sequentielle Ereignisanalyse	Auswahl von Personen mit Kundenkontakt zum Dienstleister	hoher Analyseaufwand (offene Fragen; Leitfadeninterviews)	aufgrund der aufwändigen Erhebungsmethodik oft nur kleine Stichproben realisierbar	Maßnahmenentwicklung	gegeben
Critical Incident-Technik	Auswahl von Personen mit Kundenkontakt zum Dienstleister	hoher Analyseaufwand (offene Fragen; Leitfadeninterviews)	aufgrund der aufwändigen Erhebungsmethodik oft nur kleine Stichproben realisierbar; dauerhaft Unzufriedenheit erzeugende Umstände (z.B. Preis) werden eher nicht berichtet	Maßnahmenentwicklung	gegeben
Frequenz-Relevanz-Analyse von Problemen	Auswahl von Personen mit Kundenkontakt zum Dienstleister	relative einfache Auswertung; Aufwand durch Erzeugung einer möglichst vollständigen Liste von Problemen	Vollständigkeit der Problemliste als Voraussetzung	Maßnahmenentwicklung	gegeben
Conjoint Measurement	Auswahl (potenzieller) Kunden im Rahmen einer Haushaltsbefragung	relativ hoher Aufwand bei der Fragebogenkonstruktion (Festlegung der Eigenschaften und deren Ausprägungen)	es können nur wenige Merkmale simultan betrachtet werden	Entscheidung zwischen verschiedenen Alternativen bzgl. Angebotsgestaltung	eingeschränkt gegeben
Multiattributive Verfahren	Auswahl (potenzieller) Kunden im Rahmen einer Haushaltsbefragung oder Auswahl von Fahrgästen aus Wagenfahrten (mehrstufig)	falls Befragung / Auswahl im Fahrzeug: hoher Aufwand durch mehrstufiges Auswahlverfahren; relativ einfache Auswertung	wenig detaillierte Ergebnisse zur Maßnahmenableitung (z.B. zu den Ursachen; i.d.R. müssen Zusatzinformationen abgefragt werden)	Benchmarking; Zeitreihen der Kundenzufriedenheit; in gewissem Umfang Evaluation von Maßnahmen	eingeschränkt gegeben
Penalty-Reward-Faktoren-Analyse	(kein Erhebungs-, sondern ein Datenanalyseverfahren)	Auswertung durch spezielle Regressionstechniken	keine detaillierten Informationen z.B. zu Ursachen der (Un-)Zufriedenheit (vgl. multiattributive Verf.)	Identifikation von „Basisfaktoren“	eingeschränkt gegeben

5 Praxistipps

In diesem Kapitel sollen noch einige Hinweise zu Aspekten gegeben werden, die in der Praxis häufig nicht adäquat oder zu spät bedacht werden. Dies kann möglicherweise dazu beitragen, einige typische Fehler zu vermeiden oder zumindest zu minimieren.

Durch Befragungen lassen sich erinnerte und/oder beabsichtigte Verkehrsaktivitäten und Verkehrsverhaltenshintergründe von Personen auf der Basis der realen zeit-, raum- und soziodemografischen Bezüge erfassen. In dem nachfolgenden Schema sind die generellen Arbeitsschritte bzw. die unterschiedlichen Phasen einer Befragung dargestellt:

Abbildung 10: Ablauf einer Kundenzufriedenheitsbefragung



Die Qualität von Erhebungsergebnissen steht immer im Spannungsfeld der finanziellen und personellen Ressourcen sowie des Erhebungsumfangs und -aufwands. Die Konzeption der Erhebung resultiert demnach aus einem Abwägungsprozess zwischen der gewünschten Datenqualität, dem Erhebungsumfang und den vorhandenen finanziellen und personellen Ressourcen. Dabei ist es die Aufgabe für die Verantwortlichen der Erhebung, die in Abhängigkeit der Aufgabenstellung optimale Mischung zwischen diesen drei Elementen zu finden. In der Regel werden die finanziellen Mittel vorgegeben, so dass nur noch die Priorität zwischen Quantität und Qualität gesetzt werden kann. Methoden und Maßnahmen zur Qualitätssicherung binden i. Allg. Ressourcen, die auf der anderen Seite zu Einsparungen an der Quantität, d. h. beim Stichprobenumfang und Erhebungsinhalt je Teilnehmer führen müssen. Andererseits erhöht eine Reduktion des Stichprobenumfangs den Stichprobenfehler, so dass im Extremfall die Repräsentativität der Erhebung, d.h. die Verallgemeinerungsfähigkeit der Erhebungsergebnisse nicht mehr gegeben ist.

Die Ziele und Vorgaben zur Qualität sollten daher immer im Zusammenhang mit den zur Verfügung stehenden finanziellen Mitteln und dem geplanten Erhebungsumfang gesehen werden. Je nach Aufgabenstellung sind bei der Erhebungskonzeption entsprechende Prioritäten zu setzen. Das optimale Erhebungskonzept muss damit nicht das Konzept sein, welches zur höchsten Datenqualität führt, sondern dasjenige, welches unter den gegebenen Rahmenbedingungen, insbesondere den finanziellen Restriktionen, die Untersuchungsziele am besten erfüllt.

Unabhängig davon ist jedoch sicherzustellen, dass der gesamte Erhebungsprozess dokumentiert wird. Dies ist die Grundvoraussetzung, um die Qualität der erhobenen Daten überhaupt beurteilen zu können (siehe Bäumer et al. 2010).

Auf einen immer wieder auftretenden und unter Datenqualitätsaspekten sehr relevanten Punkt soll an dieser Stelle explizit hingewiesen werden:

In der Praxis werden Erhebungen häufig zwar mit einem ganz bestimmten Ziel durchgeführt, anschließend kommt jedoch nicht selten der Wunsch auf, die vorliegenden Daten auch für andere Zwecke einzusetzen. Dies ist im Sinne der Generierung eines „Mehrwertes“ grundsätzlich zu begrüßen, unter Datenqualitätsaspekten kann ein solches Vorgehen aber fragwürdig sein. Die Sekundärnutzung von Daten sollte sorgfältig geprüft werden, da aufgrund einer zumeist starken Ausrichtung des Erhebungskonzeptes auf bestimmte zu ermittelnde Zielgrößen sich bei veränderter Fragestellung häufig keine validen und statistisch abgesicherten Ergebnisse mehr erzielen lassen. Deshalb ist im Sinne der Qualitätssicherung darauf zu achten, dass die erhobenen Daten möglichst nur für den Zweck verwendet werden, für den sie erhoben worden sind.

Zum Abschluss noch einige generelle Hinweise:

1. Es ist wichtig, im Vorfeld genau festzulegen, was das Ziel der Kundenzufriedenheitsbefragung sein soll.
2. Es sollten bereits im Planungsstadium der Kundenzufriedenheitsbefragung alle relevanten Abteilungen im Unternehmen einbezogen werden.
3. Es kann sinnvoll sein, sich ggf. extern beraten zu lassen. Dies betrifft z.B. auch die Frage, wie eine Ausschreibung ausgestaltet sein sollte, damit die definierten Ziele auch erreicht werden können (notwendiger Stichprobenumfang etc.). Darüber hinaus ist vor allem beim Fehlen entsprechender Vorerfahrungen auch eine Beratung im Hinblick auf die rechtlichen Rahmenbedingungen und Voraussetzungen einer Befragung (z.B. Datenschutz) in Erwägung zu ziehen.
4. Untersuchungen zur Kundenzufriedenheit sollten nicht mit Werbeaktivitäten (z.B. Direktmarketing) verknüpft werden. Hier ist eine saubere Trennung zwischen Forschung und Werbung anzuraten.
5. Bei der Budgetfestlegung der Befragung sollte man sich vorab über einen realistischen Rahmen informieren. Das maximal zur Verfügung stehende Budget und die damit erreichbaren Ziele sollten frühzeitig in Einklang gebracht werden. Das veranschlagte Budget sollte rechtzeitig bei der zukünftigen Haushaltsplanung berücksichtigt werden.
6. Es sollten immer mehrere Angebote eingeholt werden, ggf. über eine begrenzte Ausschreibung. Bei der Auswahl des Dienstleisters sollten bestimmte Qualitätskriterien beachtet werden (z.B. das Zur-Verfügung-Stellen von Informationen über die bisherigen Auftraggeber des Dienstleisters mittels Referenzlisten und die Berücksichtigung national und international anerkannter Standards, wie beispielsweise dem ICC/ESOMAR-Kodex). Ebenfalls hilfreich (und dennoch sehr häufig aus unterschiedlichsten Gründen nicht vorgenommen) kann z.B. eine telefonische Nachfrage bei anderen Auftraggebern sein, die in letzter Zeit vergleichbare Ausschreibungen durchgeführt haben.
7. Im Falle der Einbeziehung externer Dienstleister sollte nach der Vergabe der Befragung ein Abstimmungsgespräch mit dem Auftragnehmer stattfinden. Dabei sollten vom Auftraggeber die Ziele und der inhaltliche Rahmen vorgegeben werden, aber der Wunsch nach Mitgestaltung hinsichtlich jeder Fragenformulierung zurückgeschraubt werden, um die Vorbereitung

der Befragung nicht unnötig in die Länge zu ziehen, denn schließlich wurde der Auftragnehmer ja als Experte genau für diese Aufgabe ausgesucht.

8. Regelmäßige Abstimmungen sind genauso wichtig, wie sich die Ergebnisse präsentieren und erläutern zu lassen. Der Auftragnehmer sollte außerdem nach daraus resultierenden konkreten Umsetzungsvorschlägen gefragt werden, sofern dies nicht sowieso Auftragsbestandteil war oder der Auftragnehmer nicht selbst bereits mit der Berichtsabgabe entsprechende Vorschläge unterbreitet hat.
9. Wichtig: Die Ergebnisse einer Marktforschung müssen dann auch zu Entscheidungen führen. Deshalb sollte bereits im Vorfeld festgelegt werden, welche Schritte im Unternehmen nach Vorlage der Ergebnisse zu erfolgen haben.
10. Abschließend sei noch auf einen Aspekt hingewiesen, der aus verschiedenen Gründen häufig vernachlässigt wird: Werden aus den Ergebnissen von Kundenzufriedenheitsuntersuchungen Maßnahmen abgeleitet, so sollte nach einer angemessenen Zeit (wenn sich die Wirkung der Maßnahme entfaltet und verstetigt hat) eine Erfolgskontrolle durchgeführt werden.

Literatur

- Ackermann, T. (2010): Marktforschung ist die Grundlage für Marketing. *Bus & Bahn* 5/2010, S.12-13
- Albrecht, K. & Bradford, L. J. (1990). *The Service Advantage - How to Identify and Fulfill Customer Needs*. Homewood: Dow Jones-Irwin.
- Angermüller, W. M. (2007). Die Kundenzufriedenheit und ihr Einfluss auf die Nachfrage. Ein strategisches Instrument zur Erlössteuerung. *Der Nahverkehr*, 7-8, S. 56 - 61.
- Bäumer, M. (2004): Evaluierung von Mobilitätsinformationsangeboten. In: *Verkehrssystem- und Mobilitätsmanagement im ländlichen Raum - Studien zur Mobilitäts- und Verkehrsforschung*, Bd. 10. Hg.: Andreas Kagermeier. Mannheim, S. 217-231.
- Bäumer, M., Hussinger, A., Pfeiffer, M. & Nickel, B.E. (2009). Kundenzufriedenheit und wirtschaftlicher Erfolg im Nahverkehr. *Der Nahverkehr*, H. 11, S. 22-27.
- Bäumer, M., Hautzinger, H., Kathmann, T., Schmitz, S., Sommer, C. & Wermuth, M. (2010). *Ermittlung von Standards für anforderungsgerechte Datenqualität bei Verkehrserhebungen*, Bundesanstalt für Straßenwesen, Reihe Verkehrstechnik, Heft V 200, Bremerhafen: NW-Verlag
- Bäumer, M., Hussinger, A. & Pfeiffer, M. (2010): *Bedeutung von Kundenzufriedenheit im ÖPNV: Kundenzufriedenheitsmodelle und Messverfahren als Grundlage für die Ableitung von Maßnahmen der Angebotsgestaltung im Marketing-Mix*. Schlussbericht zum FOPS-Projekt FE 70.796/2006. Mannheim: IVT Research.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden* (11. Auflage). Berlin: Springer.
- Bogner, W. & Mayer, M. (1994). Einsatz der Conjoint-Methode. *Der Nahverkehr*, 4, S. 63 - 66.
- Bogner, W. & Plotegher, M. (1997). Wie und womit die Fahrgäste im ÖPNV zufrieden sind. *Der Nahverkehr*, 6, S. 31 - 36.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Auflage). Berlin: Springer.

- Brög, W. & Kahn, T. (2003). *Kundenzufriedenheit als ein Element strategischer Unternehmensentscheidungen*. Beitrag zur „2. UITP Internationale Marketing Konferenz“, 12. - 14. November, Paris.
- Buschmann, V. (2000). Marktforschung mit Qualitätstestern. In Deutsche Verkehrswissenschaftliche Gesellschaft (Hrsg.), *Methoden der Marketingforschung von Mobilitätsdienstleistern. Schriftenreihe der DVWG, B 238* (S. 69 – 77). Bergisch Gladbach.
- Dahlmann-Resing, T., Gebauer, S. & Klodner, H. (2004). Messung der Dienstleistungsqualität. VAG Nürnberg koppelt die Ermittlung der Kundenzufriedenheit und der Leistungsqualität. *Der Nahverkehr*, 6, S. 31 - 35.
- Diekmann, A. (2007). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen* (18. Auflage). Reinbek: Rowohlt.
- Dillman, D.A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley-Interscience
- Dillman, D.A. (1999). *Mail and Internet Surveys. The Tailored Design Method*. New York: John Wiley & Sons
- Felscher-Suhr, U., Mattern, U. & Schreckenber, D. (2005). Testkunden für justiziable Messung der Dienstleistungsqualität. Mystery Shopping in der Region Hannover als Bestandteil des vertraglichen Qualitätssteuerungssystems. *Der Nahverkehr*, 4, S. 31 - 34.
- Forschungsgesellschaft für Straßen- und Verkehrswesen (2011). Neufassung der Empfehlungen für Verkehrserhebungen (EVE), Bonn
- Hartung, J., Elpelt, B., Klösener, K.-H. (2009). Statistik. Lehr- und Handbuch der angewandten Statistik, 15. überarb. und erw. Auflage. München/Wien: Oldenbourg
- Hentschel, B. (1992). *Dienstleistungsqualität aus Kundensicht. Vom merkmals- zum ereignisorientierten Ansatz*. Wiesbaden: Deutscher Universitäts-Verlag (DUV).
- Hentschel, B. (2000). Multiattributive Messung der Dienstleistungsqualität. In M. Bruhn, M. & B. Stauss (Hrsg.), *Dienstleistungsqualität. Konzepte - Methoden – Erfahrungen*, 3. Auflage (S. 289 – 320). Wiesbaden: Gabler.

- Herr, C. (1998). Der Fahrgastbeirat im VBN. Ein Beitrag zur Kundenorientierung im ÖPNV in Bremen/Niedersachsen. *Der Nahverkehr*, 5, S. 71ff.
- Herrmann, A., Bauer, H. H. & Herrmann, S. (1996). Kundenorientierte Gestaltung des öffentlichen Personennahverkehrs. *Zeitschrift für Verkehrswissenschaft*, 4, S. 327 - 337.
- Herrmann, A. (1998). Das Kundenzufriedenheitskonzept - Ein Ansatz für Betriebe des öffentlichen Personennahverkehrs. *Zeitschrift für Verkehrswissenschaft*, 2, S. 130 - 143.
- Homburg, C. & Rudolph, B. (1997). Theoretische Perspektiven zur Kundenzufriedenheit. In H. Simon & C. Homburg (Hrsg.), *Kundenzufriedenheit. Konzepte - Methoden – Erfahrungen*, 2. Auflage (S. 31 – 51). Wiesbaden: Gabler.
- Homburg, C. & Stock-Homburg, R. (2006). Theoretische Perspektiven zur Kundenzufriedenheit. In C. Homburg (Hrsg.), *Kundenzufriedenheit. Konzepte - Methoden – Erfahrungen*, 6. Auflage (S. 17 – 51). Wiesbaden: Gabler.
- Homburg, C. & Fürst, A. (2008). Überblick über die Messung von Kundenzufriedenheit und Kundenbindung. In M. Bruhn, & C. Homburg (Hrsg.), *Handbuch Kundenbindungsmanagement. Strategien und Instrumente für ein erfolgreiches CRM*, 6. Auflage (S. 607 – 642). Wiesbaden.
- Höfler, L. & Czermak, P. (2004). Feigenblatt oder wirksames Instrument für Fahrgastnutzen? Der Benutzerbeirat für Nahverkehrplanungen im Großraum Linz: Eine Zwischenbilanz. *Der Nahverkehr*, 3, S. 59ff.
- Isfort, A. (2000). Das ÖPNV-Kundenbarometer. Instrument zur Standortbestimmung für kleine und mittlere Unternehmen und Verbände. *Der Nahverkehr*, 11, S. 39 - 42.
- Kepper, J. (2007). Ein Jahr Fünf-Minuten-Garantie. *Der Nahverkehr*, 9, S. 64 - 67.
- Krietemeyer, H. (1990). Die Conjoint-Measurement-Technik - ein neues Analyseverfahren. *Der Nahverkehr*, 4, S. 64 - 67.
- Krietemeyer, H. & Oelbaum, P. (1997). Das Kundenbarometer. Ein Instrument zur Verstärkung der Kundenorientierung. *Der Nahverkehr*, 3, S. 22 - 30.

- Kroeber-Riel, W. & Weinberg, P. (1996). *Konsumentenverhalten* (6. Auflage). München: Verlag Vahlen.
- Lindenberg, W., Prella, M., Keese, T. & Uhlenhut, A. (2007). Üstra nimmt Qualitätstests erfolgreich selbst in die Hand. Vorteile gegenüber Mystery Shopping Dritter. *Der Nahverkehr* 3, S. 54 - 59.
- Lüttenberg, R., Person, D. & Felscher-Suhr, U. (2004). Dienstleistungsqualität gezielt ermitteln und verbessern. Ergebnisse eines Mystery Shoppings bei der Straßenbahn Herne-Castrop-Rauxel. *Der Nahverkehr*, 7-8, S. 56ff.
- Mattern, U., Zielstorff, H. & Raabe, T. (2001). Nachfragepotentiale mit Mobilitätspaketen ausschöpfen. Limit-Conjoint-Analyse von ÖPNV-Nutzern in der Region Hannover. *Der Nahverkehr* 10, S. 7 - 11.
- Meffert, H. & Bruhn, M. (2008). *Dienstleistungsmarketing. Grundlagen - Konzepte - Methoden* (6. Auflage). Wiesbaden: Gabler.
- Meyer, A. & Dornach, F. (1995). Nationale Barometer zur Messung von Qualität und Kundenzufriedenheit bei Dienstleistungen. In M. Bruhn & B. Stauss (Hrsg.), *Dienstleistungsqualität. Konzepte - Methoden – Erfahrungen*, 2. Auflage, S. 429 - 453. Wiesbaden: Gabler.
- Parasuraman, A., Zeithaml, V. A. & Berry, L. L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 1, S. 12 - 40.
- Parasuraman, A., Berry, L. L. & Zeithaml, V. A. (1991). Understanding Customer Expectations of Service. *Sloan Management Review*, 3, S. 39 - 48.
- Pfeiffer, M. (2001): Zufriedenheitsmessung durch Kundenbefragungen. Ereignis- oder merkmalsorientierte Messung? *Der Nahverkehr*, 1+2, S. 48-52.
- Pietsch, W., Lemmermann, J. & Brunjes, H. (1994). Bremer Straßenbahn beschreitet mit Fahrgastforum offensiven Kurs für die Zukunft des ÖPNV. *Bus & Bahn*, 9, S. 12 - 13.
- Porst, R. (2000). Question Wording - Zur Formulierung von Fragebogen-Fragen. *ZUMA How-to-Reihe Nr. 2*. Mannheim

- Probst, G., Kahrs, C. & Rieschick-Dziabas, S. (2001). Was bindet die Fahrgäste dauerhaft? Ereignisorientierte Kundenzufriedenheit im ÖPNV. *Der Nahverkehr*, 3, S. 8 - 12.
- Richardson, A.J., Ampt, E.S. & Meyburg, A.H. (1995). *Survey methods for transport planning*. Melbourne: Eucalyptus Press
- Rosenberg, M. J. (1956). Cognitive Structure and Attitudinal Affect. *Journal of Abnormal and Social Psychology*, 11, S. 367 - 372.
- Schiefelbusch, M. & Fliegel, M. (2006). Fahrgastbeiräte als Ratgeber für Verkehrsunternehmen. Konzepte und Erfahrungen in Deutschland und Österreich. *Der Nahverkehr*, 7-8, S. 48 - 53.
- Schnippe, C. (1999). Relevanz von Qualitätskriterien. *Der Nahverkehr*, 4, S. 52ff.
- Schnippe, C. (1998). Die Messung der Kundenzufriedenheit: Praktische und psychologische Hinweise für Verkehrsunternehmen. *Der Nahverkehr*, 3, S. 32 - 38.
- Schnippe, C. (2000). *Psychologische Aspekte der Kundenorientierung. Die Kundenzufriedenheit mit der Qualität von Dienstleistungsinteraktionen am Beispiel des ÖPNV*. Frankfurt am Main: Lang.
- Schulz, M. & Meinhold, C. (2003). Quantifizierung des Schienenbonus. Messung des Kundennutzens mittels Choice-Based-Conjoint-Analyse. *Der Nahverkehr*, 6, S. 26 - 29.
- Stauss, B. (2000). „Augenblicke der Wahrheit“ in der Dienstleistungserstellung. Ihre Relevanz und ihre Messung mit Hilfe der Kontaktpunkt-Analyse. In M. Bruhn & B. Stauss (Hrsg.), *Dienstleistungsqualität. Konzepte - Methoden – Erfahrungen*, 3. Auflage, S. 321 - 340. Wiesbaden: Gabler.
- Stauss, B. & Seidel, W. (2006). Prozessuale Zufriedenheitsermittlung und Zufriedenheitsdynamik bei Dienstleistungen. In C. Homburg (Hrsg.), *Kundenzufriedenheit. Konzepte - Methoden – Erfahrungen*, 6. Auflage, S. 171 - 195. Wiesbaden: Gabler.
- Strasser, S., Haller, M. & Krietemeyer, H. (2008). Kundenpotenziale ausschöpfen. ÖPNV-Marktanteile und Zielgruppen für das Marketing im Großraum München. *Der Nahverkehr*, 1-2, S. 51 - 59.

-
- Trommsdorff, V. (1975). *Die Messung von Produktimages für das Marketing. Grundlagen und Operationalisierung*. Köln: Heymanns.
- Vetter, T. & Haase, K. (2008). Alternative Bedienformen im ÖPNV – Akzeptanzstudie im Landkreis Saalkreis. Institut für Wirtschaft und Verkehr (Hrsg.) *Diskussionsbeiträge aus dem Institut für Wirtschaft und Verkehr*, 1. Verfügbar unter: http://tu-dresden.de/die_tu_dresden/fakultaeten/vkw/iwv/diskuss/diskbeitrag_2008_1.pdf [28.01.2010]
- Verband Deutscher Verkehrsunternehmen (VDV) (1994). *Jahresbericht 1993*. Köln: Moecker Merkur.
- Verband Deutscher Verkehrsunternehmen (VDV) (2002). Messung der Dienstleistungsqualität im ÖPNV. Methodenbewertung unter dem Aspekt von Bonus-/Malus-Regelungen. *VDV-Mitteilungen 10008*. Köln.
- Verband Deutscher Verkehrsunternehmen (VDV) (2005). Marktforschung im ÖPNV. *VDV-Mitteilungen 9022*. Köln.
- Verband Deutscher Verkehrsunternehmen (VDV) (2010). Gewinnung und Verknüpfung von Nachfragedaten im ÖPNV. *VDV-Schriften 951*. Köln.
- Wagener, N. & Jahn, D. (2000). Qualität im Wettbewerb. Kundenzufriedenheitsforschung im liberalisierten ÖPNV-Markt. *Der Nahverkehr*, 4, S. 67 - 70.
- Wiethüchter, J., Eidam, H., Knoch, C. & Dallmeyer, H. (2007). Im Dialog mit den Kunden. Internetgestütztes Beschwerdemanagement beim Stadtverkehr Hürth. *Der Nahverkehr*, 4, S. 45-48.
- Zumkeller, D., Blechinger, W., Chlond, B. und Seitz, H. (1994). *Paneluntersuchungen zum Verkehrsverhalten*. Forschung Straßenbau und Straßenverkehrstechnik, Heft 688. Bonn