IWD
market research

FLUGGASTBEFRAGUNGEN AM FLUGHAFEN

Wie Passagierwahrnehmung messbar und steuerbar wird.
Ein praxisbasiertes Methodenwhitepaper zu Befragungen im Luftverkehr.

Autorinnen: Kordula Later, Sandra Baethge

WHITEPAPER 3.0 / STAND 10.2026

Fluggastbefragungen am Flughafen

Inhalt

Executive Summary

1. Ausgangslage: Was operative Kennzahlen zeigen und was nicht

2. Anwendungsfelder am Flughafen

3. Der Flughafen als Forschungsfeld
   - 3.1 Was Marktforschung am Flughafen besonders macht
   - 3.2 Zwei Studientypen: standardisierte Erhebungen und standortspezifische Studien

4. Methodische Grundlagen
   - 4.1 Erhebungsformen: CAPI, CASI, Mixed-Mode und Online
   - 4.2 Stichprobe, Quotierung und Auswahl der Passagiere
   - 4.3 Fallzahl und Genauigkeit
   - 4.4 Fragebogen: standardisierter Kern, gezielte Module
   - 4.5 Feldarbeit und Qualitätssicherung

5. Kennzahlen und Auswertung
   - 5.1 Kennzahlen-Systematik: Struktur, Erleben, Wirkung
   - 5.2 Von Daten zu Erkenntnissen

6. Methodische Herausforderungen und Limitationen
   - 6.1 Zeitdruck und Befragungssituation
   - 6.2 Repräsentativität und seltene Situationen
   - 6.3 Internationale Passagiere

7. Fazit: Von der Befragung zur Entscheidung

Quellen

Executive Summary

Flughäfen gehören zu den am genauesten vermessenen Betrieben überhaupt. Ob Durchlaufzeiten an der Sicherheitskontrolle, Auslastung der Check-in-Schalter, Belegung der Parkhäuser oder Umsätze je Passagier, nahezu jeder Prozessschritt erzeugt Daten, oft in Echtzeit. Diese Daten zeigen zuverlässig, was am Flughafen geschieht. Sie zeigen jedoch nicht, wie Passagiere es erleben.

Genau diese Lücke ist steuerungsrelevant. Ob sich Passagiere sicher orientieren, ob sie eine Wartezeit als zumutbar empfinden, warum sie ein Gastronomieangebot meiden oder einen Flughafen weiterempfehlen, lässt sich aus den laufenden Kennzahlen eines Flughafens nicht ablesen. Fluggastbefragungen schließen diese Lücke. Sie verbinden das Erleben der Passagiere mit den operativen Kennzahlen und machen Ursachen sichtbar, wo Kennzahlen nur Symptome zeigen.

Dieses Whitepaper zeigt zunächst, in welchen Anwendungsfeldern Befragungen am Flughafen Erkenntnisse liefern, und veranschaulicht dies mit Use Cases aus typischen Fragestellungen. Anschließend erläutert es, welche Besonderheiten der Flughafen als Forschungsfeld mit sich bringt, wie Fluggastbefragungen methodisch aufgebaut sind, welche Kennzahlen sich erheben lassen und wie aus Daten belastbare Entscheidungen werden.

Zentrale Erkenntnisse dieses Whitepapers:

- Die Differenz zwischen gemessener und erlebter Leistung ist selbst eine Steuerungsgröße. Sie erklärt, warum stabile Kennzahlen und sinkende Zufriedenheit gleichzeitig auftreten können.
- Standardisierte Erhebungen wie die ADV-Fluggastbefragung oder ACI ASQ und standortspezifische Studien beantworten unterschiedliche Fragen und folgen unterschiedlichen methodischen Regeln. Die Wahl des Studientyps ist die erste und wichtigste Weichenstellung.
- Die Aussagekraft einer Befragung entscheidet sich an der Stichprobe, nicht an der Zahl der Interviews. Eine große Fallzahl schützt nicht vor einer schlechten Stichprobe.
- Der Erhebungsweg beeinflusst das Ergebnis: Ob ein Interviewer fragt oder der Passagier selbst am Tablet antwortet, verändert die Bewertungen. Wer Verfahren kombiniert oder wechselt, muss diese sogenannten Moduseffekte kontrollieren.
- Den größten Erkenntniswert liefern Befragungen, wenn sie mit den operativen Kennzahlen des Flughafens verknüpft werden und wenn die Auswertung zeigt, welche Stationen der Reise das Gesamturteil tatsächlich prägen.
- Der Flughafen ist ein anspruchsvolles Forschungsfeld. Sicherheitsauflagen, Zeitdruck, internationale Passagiere und ein 24-Stunden-Betrieb erfordern spezifische Feldkompetenz.

1. Ausgangslage:
Was operative Kennzahlen zeigen und was nicht

Kaum ein anderer Ort des öffentlichen Lebens ist so dicht mit Messpunkten ausgestattet wie ein Flughafen. Zählsysteme erfassen, wie viele Passagiere in welchem Zeitfenster die Sicherheitskontrolle passieren. Warteschlangensensoren messen Durchlaufzeiten. Kassensysteme dokumentieren Umsätze je Fläche & Destination, Parkleitsysteme die Belegung jedes Parkdecks. Beschwerdemanagement und Online-Bewertungen liefern zusätzlich ein Stimmungsbild. Zusammen bilden sie die operativen Kennzahlen, die ein Flughafen laufend in seinen Systemen und Dashboards vor sich hat. Sie sind unverzichtbar, denn ohne sie ließe sich ein Flughafen nicht steuern.

Alle diese Kennzahlen beschreiben jedoch Abläufe, nicht Erfahrungen. Eine Durchlaufzeit von zwölf Minuten kann als zügig oder als belastend erlebt werden, je nachdem, ob Passagiere wissen, wie lange sie noch warten müssen, ob der Abflug knapp bevorsteht und ob der Wartebereich übersichtlich ist. Ein Umsatzrückgang in einem Gastronomiebereich zeigt, dass weniger gekauft wird, aber nicht, ob das Angebot, der Preis, die Lage oder die verfügbare Zeit der Grund ist.

Die Lücke zwischen Messwert und Erleben

Dass erlebte und tatsächliche Wartezeit auseinanderfallen, ist in der Dienstleistungsforschung gut beschrieben. Grundlegend ist die Arbeit von Maister (1985) zur Psychologie des Wartens, deren Grundsätze bis heute im Servicemanagement angewendet werden. Danach wirken unbestimmte Wartezeiten länger als bekannte, unerklärte länger als erklärte, Wartezeiten ohne Beschäftigung länger als solche, in denen etwas geschieht, und Anspannung, etwa die Sorge, den Flug zu verpassen, lässt das Warten zusätzlich länger erscheinen. Zwei Prozesse mit identischer Durchlaufzeit können am Flughafen deshalb völlig unterschiedlich bewertet werden. Umgekehrt kann eine Maßnahme, die die Durchlaufzeit gar nicht verändert, die Zufriedenheit deutlich verbessern.

Abb.: Die Wahrnehmungsdifferenz zwischen gemessener und erlebter Wartezeit (schematische Darstellung)

Diese Differenz zwischen Messwert und Erleben, die wir in diesem Whitepaper als Wahrnehmungsdifferenz bezeichnen, ist deshalb keine Störgröße, die man herausrechnen müsste. Sie ist selbst eine steuerungsrelevante Information. Sie zeigt, wo operative Verbesserungen bei den Passagieren ankommen und wo nicht, und sie weist auf Hebel hin, die in keiner operativen Kennzahl auftauchen, etwa Information, Orientierung, Aufenthaltsqualität und das Gefühl, die Situation im Griff zu haben.

USE CASE: Wenn eine stabile Kennzahl ein sinkendes Erleben verdeckt

Ausgangslage: An einem Flughafen lag die durchschnittliche Wartezeit an der Sicherheitskontrolle über mehrere Quartale stabil im Zielkorridor. Gleichzeitig sank die Zufriedenheit mit dem Kontrollprozess in den begleitenden Passagierbefragungen von Welle zu Welle.

Was die Befragung zeigte: Die Passagiere schätzten ihre Wartezeit deutlich länger ein, als sie tatsächlich war. Die Differenz war dort am größten, wo die Warteschlange nicht einsehbar war und keine Information zur verbleibenden Wartedauer vorlag. Belastend war nicht die Wartezeit selbst, sondern die Ungewissheit.

Konsequenz: Der Flughafen installierte Anzeigen zur voraussichtlichen Wartezeit an den Zugängen zur Kontrolle. Die gemessene Durchlaufzeit blieb unverändert, die Zufriedenheit mit dem Prozess stieg in den folgenden Wellen wieder an.

Die strategische Relevanz für Flughäfen

Die Lücke zwischen Messwert und Erleben ist keine akademische Detailfrage, denn an ihr hängen wirtschaftlich relevante Entscheidungen. Ein wachsender Teil der Flughafenerlöse stammt aus dem Non-Aviation-Geschäft, also aus Retail, Gastronomie, Parken und Services. Diese Erlöse hängen unmittelbar davon ab, wie entspannt Passagiere den Flughafen erleben und wie viel Zeit ihnen nach der Sicherheitskontrolle bleibt. Wer die Aufenthaltsqualität verbessert, verbessert auch die Umsatzbasis.

Hinzu kommt der Wettbewerb um Passagiere. In vielen Regionen überschneiden sich die Einzugsgebiete mehrerer Flughäfen. Ob Reisende den nächstgelegenen Flughafen wählen oder eine längere Anfahrt in Kauf nehmen, hängt neben dem Streckenangebot auch von Erreichbarkeit, Parken und dem Gesamteindruck ab. Gegenüber Airlines wiederum müssen Flughäfen belegen können, welches Nachfragepotenzial ihr Einzugsgebiet für eine neue Verbindung bietet. Und schließlich stehen Flughäfen regelmäßig vor Investitionsentscheidungen in Millionenhöhe, vom Terminalumbau über neue Sicherheitslinien bis zum Wegeleitsystem, deren Wirkung auf die Passagiere vorab abgeschätzt und danach belegt werden sollte.

Fluggastbefragungen sind die richtige Methode, wenn:

- operative Kennzahlen unauffällig sind, Beschwerden, Online-Bewertungen oder Rückmeldungen von Mitarbeitenden aber auf Unzufriedenheit hindeuten und die Ursache unklar ist
- Investitionen in Prozesse, Flächen oder Leitsysteme vorbereitet oder in ihrer Wirkung belegt werden müssen
- die Struktur der Passagiere, ihr Einzugsgebiet und ihre Reisemotive für Marketing, Anbindung oder Airline-Akquise gebraucht werden
- Non-Aviation-Angebote hinter den Erwartungen bleiben und die Gründe dafür bei Nutzern und Nichtnutzern liegen
- die eigene Servicequalität im Vergleich zu anderen Flughäfen eingeordnet werden soll

2. Anwendungsfelder am Flughafen

Bevor es um Methodik geht, lohnt ein Blick auf die Fragen, die Flughäfen mit Befragungen beantworten. Sie reichen von der strategischen Frage, woher die Passagiere kommen und welches Potenzial im Einzugsgebiet liegt, bis zur operativen Frage, warum eine Warteschlange als belastend erlebt wird, obwohl sie sich zügig bewegt. Nahezu jeder Bereich eines Flughafens, von der Zufahrt bis zum Gate, kann Gegenstand einer Befragung sein.

Abb.: Sechs Anwendungsfelder für Befragungen am Flughafen

Die folgenden Anwendungsfelder zeigen jeweils die typische Fragestellung, was eine Befragung dazu liefert und was Flughäfen mit den Ergebnissen tun können. Use Cases veranschaulichen, wie das in der Praxis aussieht. Die methodischen Begriffe, die dabei anklingen, etwa die Unterscheidung zwischen standardisierten Erhebungen und standortspezifischen Studien, werden in den Kapiteln 3 bis 5 erläutert.

Typische Auslöser für eine Fluggastbefragung

- Ein Terminalumbau, eine neue Sicherheitslinie oder ein neues Wegeleitsystem steht an oder ist abgeschlossen
- Ein neues Retail- oder Gastronomiekonzept bleibt hinter den Erwartungen
- Beschwerden oder Online-Bewertungen deuten auf Unzufriedenheit hin, ohne dass die operativen Kennzahlen eine Ursache zeigen
- Für Airline-Gespräche oder die Planung der Verkehrsanbindung werden belastbare Daten zum Einzugsgebiet gebraucht

Einzugsgebiet, Anbindung und Streckenentwicklung

Die Frage: Woher kommen unsere Passagiere, und welches Potenzial liegt in unserem Einzugsgebiet?

Was die Befragung liefert:

Herkunft nach Postleitzahl, Reiseanlass, Reiseziel und Anreiseverkehrsmittel, hochgerechnet auf das reale Passagieraufkommen. In der Zeitreihe zeigen sich Verschiebungen im Einzugsgebiet früh. Verknüpft mit dem Reiseziel wird sichtbar, welche Regionen welche Verbindungen tragen.

Was Flughäfen damit tun können: Regionales Marketing gezielt ausrichten, Bus- und Bahnanbindungen begründen, Airlines das Nachfragepotenzial für neue Verbindungen belegen. Grundlage sind in der Regel standardisierte Erhebungen wie die ADV-Fluggastbefragung.

USE CASE: Das Einzugsgebiet verschiebt sich

Ausgangsfrage: Ein Flughafen wollte wissen, ob sich die Herkunft seiner Passagiere verändert, um regionales Marketing und Verkehrsanbindung gezielter auszurichten und in Gesprächen mit Airlines das Potenzial seines Einzugsgebiets belegen zu können.

Studiendesign: Standardisierte Strukturerhebung über mehrere Jahre, mit Quoten nach Zielflughafen, Airline und Zeitfenster, Einzugsgebiet auf Ebene der Postleitzahl sowie Gewichtung und Hochrechnung auf das reale Passagieraufkommen.

Methodische Besonderheit: Verschiebungen im Einzugsgebiet verlaufen langsam und sind in einer einzelnen Welle kaum von Zufallsschwankungen zu unterscheiden. Aussagekraft entsteht erst über eine stabile Zeitreihe mit unverändertem Fragebogenkern und gleichbleibender Methodik.

Ergebnis	Kernregion	Neue Herkunftsregion
Anteil am Passagieraufkommen über mehrere Jahre	stabil	kontinuierlich steigend
Anreise mit öffentlichen Verkehrsmitteln	hoch	gering
Reiseanlass	gemischt	überwiegend privat
Nachfrage nach Reisezielen	breit verteilt	auf wenige Ziele konzentriert

Ergebnisse: Die Zeitreihe zeigte einen wachsenden Anteil von Passagieren aus einer Region, die bislang nicht als Kernquellgebiet galt. Diese Passagiere reisten fast ausschließlich mit dem eigenen Pkw an, überwiegend privat und mit deutlicher Konzentration auf wenige Urlaubsziele. Viele von ihnen gaben an, früher überwiegend einen anderen Flughafen genutzt zu haben. Aus den Passagierzahlen allein war diese Verschiebung nicht erkennbar.

Handlungsempfehlungen: Regionales Marketing in der neuen Herkunftsregion verstärken, Parkangebote auf die wachsende Pkw-Anreise ausrichten und eine Busanbindung prüfen. Die Daten zur Zielnachfrage flossen in die Gespräche mit Airlines ein.

Limitation: Eine Befragung am eigenen Flughafen erfasst nur die Passagiere, die bereits kommen. Wie viele Reisende aus dem Einzugsgebiet einen anderen Flughafen nutzen und warum, zeigt sie nicht. Für diese Frage ist eine ergänzende Befragung der Bevölkerung im Einzugsgebiet nötig.

Prozesse an Check-in und Sicherheitskontrolle

Die Frage: Wie erleben Passagiere unsere Kernprozesse, und warum weicht ihr Urteil von den gemessenen Durchlaufzeiten ab?

Was die Befragung liefert:

Erlebte Wartezeit, die Differenz zur gemessenen Wartezeit, Stressempfinden und die Bewertung der einzelnen Prozessschritte, verknüpft mit den gemessenen Durchlaufzeiten desselben Zeitfensters. Befragt wird unmittelbar nach dem Prozess, solange die Erinnerung frisch ist.

Was Flughäfen damit tun können: Maßnahmen jenseits der Durchlaufzeit identifizieren, etwa Wartezeitanzeigen, eine klare Warteschlangenführung oder eine bessere Vorbereitung der Passagiere auf die Kontrolle.

Aus der Praxis: An zwei Check-in-Bereichen desselben Terminals war die Abfertigungsdauer nahezu identisch, das Stressempfinden aber deutlich verschieden. An einem Bereich gab es klar abgegrenzte Wartespuren, am anderen keine erkennbare Schlangenstruktur, und Passagiere waren unsicher, ob sie richtig anstanden. Die Befragung zeigte, dass die Ordnung des Wartens mehr zählte als seine Dauer.

Orientierung und Wegeführung

Die Frage: Finden Passagiere ihren Weg, und an welchen Entscheidungspunkten entsteht Unsicherheit?

Was die Befragung liefert:

Unsicherheit an konkreten Wegeabschnitten, genutzte Informationsquellen und Umwege. Besonders aussagekräftig sind Erstnutzer und Passagiere mit ungewohnten Wegen. Vielflieger kennen den Weg und bewerten die Beschilderung positiv, weil sie sie nicht mehr brauchen.

Was Flughäfen damit tun können: Wegweiser an den tatsächlichen Entscheidungspunkten platzieren, Online-Information und Beschilderung vor Ort aufeinander abstimmen, Veränderungen im Vorher-Nachher-Vergleich belegen. Eye-Tracking kann ergänzend zeigen, welche Wegweiser überhaupt wahrgenommen werden.

Aus der Praxis: Belegungsdaten zeigten, dass ein Parkhaus kaum genutzt wurde, während andere ausgelastet waren. Die Befragung ergab, dass viele Reisende es nicht kannten oder die Zufahrt zu spät erkannten und sich auf dem Weg vom Parkhaus zum Terminal unsicher fühlten. Zufahrtsbeschilderung und Online-Information wurden daraufhin angepasst.

USE CASE: Die Wirkung eines Terminalumbaus belegen

Ausgangsfrage: Ein Flughafen baute seinen Abflugbereich um, mit neuer Wegeführung zu den Gates, zentralisierter Sicherheitskontrolle und erweiterten Aufenthaltsflächen. Die Wirkung der Investition auf die Passagiere sollte belegt und Nachsteuerungsbedarf früh erkannt werden.

Studiendesign: Standortspezifische Studie als Vorher-Nachher-Messung in drei Wellen vor dem Umbau, kurz nach der Eröffnung und nach einigen Monaten, jeweils mit identischem Fragebogen, gleichen Befragungsorten und vergleichbaren Zeitfenstern. Ergänzend wurden Frequenzdaten der neuen Wege genutzt.

Methodische Besonderheit: Eine Veränderung lässt sich nur belegen, wenn alles andere gleich bleibt. Zudem reagieren Passagiergruppen unterschiedlich. Erstnutzer erleben den neuen Zustand unvoreingenommen, Vielflieger vergleichen ihn mit ihren gewohnten Wegen. Beide Gruppen wurden deshalb getrennt ausgewertet.

Ergebnis	Vor dem Umbau	Kurz nach Eröffnung	Nach einigen Monaten
Orientierungssicherheit Erstnutzer	mittel	deutlich höher	deutlich höher
Orientierungssicherheit Vielflieger	hoch	niedriger	wieder hoch
Aufenthaltsqualität	mittel	höher	höher
Erlebte Wartezeit Sicherheitskontrolle	hoch	niedriger	niedriger

Ergebnisse: Der Umbau wirkte. Erstnutzer orientierten sich deutlich sicherer, Aufenthaltsqualität und erlebte Wartezeit verbesserten sich. Kurz nach der Eröffnung fühlten sich jedoch ausgerechnet die Vielflieger unsicherer, weil ihre gewohnten Wege nicht mehr funktionierten. In einer Gesamtauswertung wäre dieser Effekt im positiven Durchschnitt untergegangen.

Handlungsempfehlungen: Übergangsweise zusätzliche Hinweise an den früheren Entscheidungspunkten. Die dritte Welle bestätigte, dass sich die Orientierung der Vielflieger nach einigen Monaten normalisiert hatte. Für künftige Umbauten wurde eine gezielte Information der Vielflieger vor der Eröffnung eingeplant.

Retail, Gastronomie und Non-Aviation

Die Frage: Warum nutzen Passagiere unsere Angebote nicht so, wie wir es erwarten?

Was die Befragung liefert:

Nutzungsquoten nach Passagiergruppe, Gründe für Nutzung und Nichtnutzung, verfügbarer Zeitpuffer und Preiswahrnehmung. Entscheidend ist die Befragung im Wartebereich, die Nutzer und Nichtnutzer gleichermaßen erreicht. Eine Befragung im Geschäft selbst überschätzt die Zufriedenheit systematisch.

Was Flughäfen damit tun können: Flächen- und Konzeptentscheidungen auf die Barrieren der Nichtnutzer ausrichten, Angebote auf Zielgruppen und Zeitbudgets abstimmen. Oft entscheidet der Zeitpuffer nach der Sicherheitskontrolle stärker über die Nutzung als das Angebot selbst. Ausgabenhöhen werden in Befragungen allerdings ungenau erinnert; hier sind Kassendaten die bessere Quelle, die Befragung liefert die Gründe dahinter.

USE CASE: Zufriedene Gäste, zu wenige Wiederkehrer

Ausgangsfrage: Ein neues Gastronomiekonzept im Abflugbereich wurde von seinen Gästen durchweg positiv bewertet, vor allem Qualität, Freundlichkeit und Ambiente. Der Umsatz blieb dennoch hinter der Planung zurück. Der Flughafen wollte wissen, ob Konzept, Preis oder Lage das Problem waren.

Studiendesign: Standortspezifische Studie im Wartebereich der angrenzenden Gates, nicht im Lokal selbst. Befragt wurden Nutzer und Nichtnutzer nach Zählregel, verteilt über alle Tageszeiten. Neben der Bewertung wurden Bekanntheit, frühere Besuche, Gründe für die Nichtnutzung und der verfügbare Zeitpuffer erhoben und mit den Kassendaten je Stunde verknüpft.

Methodische Besonderheit: Eine Befragung im Lokal hätte nur bestätigt, was bereits bekannt war, nämlich dass das Konzept gefällt. Die Erkenntnis lag bei denen, die nicht kamen, und bei denen, die nicht wiederkamen.

Passagiergruppe	Bewertung des Angebots	Absicht, (wieder) zu kommen	Wichtigster Grund für Zurückhaltung
Erstbesucher	hoch	mittel	Unsicherheit über die Wartezeit
Wiederholte Besucher, v. a. Vielflieger	hoch	kaum höher als bei Erstbesuchern	frühere lange Wartezeit in der Abflugspitze
Nichtnutzer	nicht bewertet	gering	Zeitpuffer zu knapp, Sorge um den Flug

Ergebnisse: Das Konzept war bekannt und geschätzt. Ein relevanter Teil der Passagiere mied es jedoch gezielt zu den Abflugspitzen, weil sie bei einem früheren Besuch eine unerwartet lange Wartezeit erlebt hatten und befürchteten, ihren Flug zu verpassen. Das Problem war nicht das Angebot, sondern die Unsicherheit über die Dauer. Gerade Vielflieger, die für Wiederholungsumsätze entscheidend sind, hatten diese Erfahrung am häufigsten gemacht.

Handlungsempfehlungen: Wartezeit an der Theke sichtbar machen und in den Spitzen verkürzen, etwa durch vorbereitete Schnellangebote oder zusätzliche Ausgabepunkte. Die Wirkung wurde in einer Folgewelle bei denselben Passagiergruppen gemessen.

Information und Störungskommunikation

Die Frage: Erreichen Informationen zu Verspätungen, Gate-Wechseln und Störungen die Passagiere rechtzeitig und verständlich?

Was die Befragung liefert:

Über welchen Kanal eine Information zuerst ankam, welchem Kanal Passagiere im Zweifel vertraut haben und ob widersprüchliche Angaben Unsicherheit erzeugt haben. Aus dem Abstand zwischen schnellstem und vertrauenswürdigstem Kanal wird sichtbar, wo Anzeigetafel, Durchsage und App nicht zusammenspielen.

Was Flughäfen damit tun können: Informationsketten abstimmen, Bezeichnungen vereinheitlichen, Gehzeiten zum neuen Gate kommunizieren. Weil Störungen selten und nicht planbar sind, braucht diese Fragestellung eine eigene Stichprobenlogik (siehe Abschnitt 6.2).

Zielgruppen: Geschäftsreisende, Familien, Umsteiger

Die Frage: Welche Passagiergruppen haben besondere Bedürfnisse, und wo bedient der Flughafen sie schlechter als andere?

Was die Befragung liefert:

Segmentierte Ergebnisse für Gruppen, die im Durchschnitt untergehen, etwa Familien mit Kindern, die an der Sicherheitskontrolle andere Probleme haben als Alleinreisende, Umsteiger ohne Ortskenntnis, Geschäftsreisende mit minimalem Zeitpuffer. Typologien verbinden Reiseanlass, Zeitbudget und Reiseerfahrung zu Gruppen mit klar unterscheidbaren Erwartungen.

Was Flughäfen damit tun können: Services zielgruppengerecht gestalten, etwa Familienspuren, Transferinformationen oder Fast-Track-Angebote, und ihre Wirkung in der jeweiligen Gruppe messen.

USE CASE: Familien an der Sicherheitskontrolle

Ausgangsfrage: Die Zufriedenheit mit der Sicherheitskontrolle lag an einem Flughafen insgesamt im guten Bereich. In den Sommerferien häuften sich jedoch Beschwerden von Familien über Hektik und Unsicherheit an der Kontrolle.

Studiendesign: Standortspezifische Studie unmittelbar nach der Sicherheitskontrolle, gezielt aufgestockt in den Ferienzeiten, mit getrennter Auswertung für Familien mit Kindern, Alleinreisende und Geschäftsreisende.

Ergebnisse: Familien bewerteten die Kontrolle deutlich schlechter als alle anderen Gruppen, obwohl ihre Wartezeit kaum länger war. Belastend waren vor allem die Vorbereitung an der Ablage, die Unsicherheit über Flüssigkeiten und Kinderwagen sowie das Gefühl, nachfolgende Passagiere aufzuhalten. Im Gesamtdurchschnitt war dieser Befund nicht sichtbar.

Handlungsempfehlungen: In Ferienzeiten eine eigene Familienspur mit mehr Platz an der Ablage einrichten und Familien schon vor der Anreise sowie im Wartebereich vor der Kontrolle gezielt informieren. Die Wirkung wurde in der folgenden Ferienzeit mit derselben Methode gemessen.

3. Der Flughafen als Forschungsfeld

3.1 Was Marktforschung am Flughafen besonders macht

Methodisch folgen Fluggastbefragungen denselben Grundregeln wie jede andere quantitative Befragung. Dazu gehören eine klar definierte Grundgesamtheit, also die Gruppe, über die eine Aussage getroffen werden soll, eine kontrollierte Stichprobe, ein getesteter Fragebogen und eine sorgfältige Auswertung. Die Bedingungen, unter denen diese Regeln umgesetzt werden müssen, sind am Flughafen jedoch besonders. Fünf Faktoren prägen jede Studie.

Sicherheitsbereich und Zugang. Die aussagekräftigsten Befragungsorte, etwa die Wartebereiche an den Gates, liegen hinter der Sicherheitskontrolle. Interviewer benötigen dafür Flughafenausweise und in Deutschland eine Zuverlässigkeitsüberprüfung nach dem Luftsicherheitsgesetz. Ausrüstung wie Tablets muss durch die Kontrolle, Aufenthaltszeiten und Bewegungsräume sind geregelt. Das erfordert einen Vorlauf von mehreren Wochen und eine enge Abstimmung mit dem Flughafen.

Enge Zeitfenster in der Reisekette. Passagiere halten sich am Flughafen nicht um seiner selbst willen auf, sondern sind auf dem Weg zu einem Flug. Vor der Sicherheitskontrolle steht die Sorge, rechtzeitig durchzukommen, im Vordergrund; kurz vor dem Boarding fehlt die Zeit. Das günstigste Fenster liegt typischerweise im Wartebereich nach der Kontrolle und vor Beginn des Boardings. Wer zu einem anderen Zeitpunkt befragen muss, braucht einen entsprechend kürzeren Fragebogen.

Heterogene und internationale Passagiere. Kaum eine Befragungssituation ist so vielfältig. Geschäftsreisende mit minimalem Zeitpuffer, Familien mit Kindern, Erstflieger, Umsteiger ohne Ortskenntnis und Passagiere aus Dutzenden Ländern treffen hier aufeinander. Sprachversionen, Ansprache und Interviewereinsatz müssen darauf ausgerichtet sein.

Betrieb rund um die Uhr und saisonale Schwankungen. Das Passagierprofil eines frühen Morgens unterscheidet sich deutlich von dem eines Nachmittags, das eines Ferienbeginns von dem einer Messewoche. Eine Studie, die nur zu bequemen Tageszeiten erhebt, bildet nur einen Ausschnitt ab.

Abhängigkeit vom Flugplan. Der Flugplan bestimmt, wann welche Passagiere an welchem Ort anzutreffen sind. Er ist damit die wichtigste Planungsgrundlage für Stichprobe, Einsatzplanung und Sprachkompetenz im Feld und wird zugleich durch Verspätungen, Gate-Wechsel und Annullierungen täglich verändert. Feldteams müssen darauf kurzfristig reagieren können.

3.2 Zwei Studientypen: standardisierte Erhebungen und standortspezifische Studien

Wer an einem Flughafen Passagiere befragen möchte, bewegt sich zwischen zwei grundlegend verschiedenen Ansätzen. Beide haben ihre Berechtigung, aber sie beantworten unterschiedliche Fragen, und die methodischen Regeln des einen lassen sich nicht ohne Weiteres auf den anderen übertragen.

Standardisierte Erhebungen werden nach einer einheitlichen, zentral festgelegten Methodik an vielen Flughäfen durchgeführt. In Deutschland ist dies vor allem die vom Flughafenverband ADV koordinierte Fluggastbefragung. Sie ist in erster Linie eine Passagierstrukturerhebung und beschreibt, wer einen Flughafen nutzt, woher die Passagiere kommen, wohin und warum sie reisen und wie sie zum Flughafen gelangen (Flughafenverband ADV 2025, o. J.). International verbreitet ist das Programm Airport Service Quality (ACI ASQ) des Airports Council International, das die Servicequalität an über 400 Flughäfen in 110 Ländern nach einheitlichem Fragebogen erhebt und vergleichbar macht (ACI World o. J.). Gemeinsam ist beiden ein standardisierter Fragebogenkern und eine Stichprobe, die aus Flugplan und Passagieraufkommen abgeleitet wird. Ihr besonderer Wert liegt in der Vergleichbarkeit zwischen Flughäfen und über die Zeit. Zu beachten ist allerdings, dass die ADV-Fluggastbefragung nur alle drei Jahre neue Daten liefert.

Standortspezifische Studien gehen von einer konkreten Fragestellung des einzelnen Flughafens aus, etwa warum ein neues Gastronomiekonzept kaum genutzt wird, wie Passagiere die umgebaute Sicherheitskontrolle erleben oder ob Umsteiger den Weg zum Anschlussflug finden. Fragebogen, Befragungsort, Zeitplanung und Stichprobe werden vollständig auf diese Frage zugeschnitten.

Merkmal	Standardisierte Erhebungen (z. B. ADV, ACI ASQ)	Standortspezifische Studien
Kernfrage	Wer nutzt den Flughafen, und wie steht er im Vergleich da?	Wie werden konkrete Prozesse, Services oder Angebote am eigenen Standort erlebt?
Fragebogen	Einheitlicher, zentral vorgegebener Kern; ergänzbar um standortspezifische Module	Vollständig auf die Fragestellung zugeschnitten
Grundgesamtheit	Alle abfliegenden Passagiere eines Zeitraums	Passagiere, die den untersuchten Bereich nutzen oder nutzen könnten
Stichprobe	Aus Flugplan und Passagierzahlen (PAX) abgeleitet, Quoten nach Zielflughafen, Airline und Zeitfenster	Aus der Fragestellung abgeleitet, Steuerung nach relevanten Situationen
Gewichtung	Anpassung an die reale Verkehrsstruktur	In der Regel keine Hochrechnung auf alle Passagiere
Stärke	Vergleichbarkeit, Zeitreihen, repräsentative Strukturdaten	Tiefe, Ursachenanalyse, unmittelbarer Bezug zu Maßnahmen

Die beiden Ansätze schließen sich nicht aus, sondern ergänzen sich. Standardisierte Erhebungen lassen sich um standortspezifische Module erweitern. Diese Module profitieren von der repräsentativen Stichprobe und lassen sich nach den erhobenen Strukturmerkmalen, etwa Reiseanlass oder Herkunft, auswerten. Nicht geeignet sind sie, wenn die untersuchte Situation zu selten vorkommt oder zu weit vom Befragungsort entfernt ist, um im Rahmen der Strukturerhebung genügend Fälle zu erreichen. Dann ist eine eigene Studie der bessere Weg. Wo sich die methodischen Regeln unterscheiden, ist in den folgenden Kapiteln vermerkt, für welchen der beiden Studientypen eine Aussage gilt.

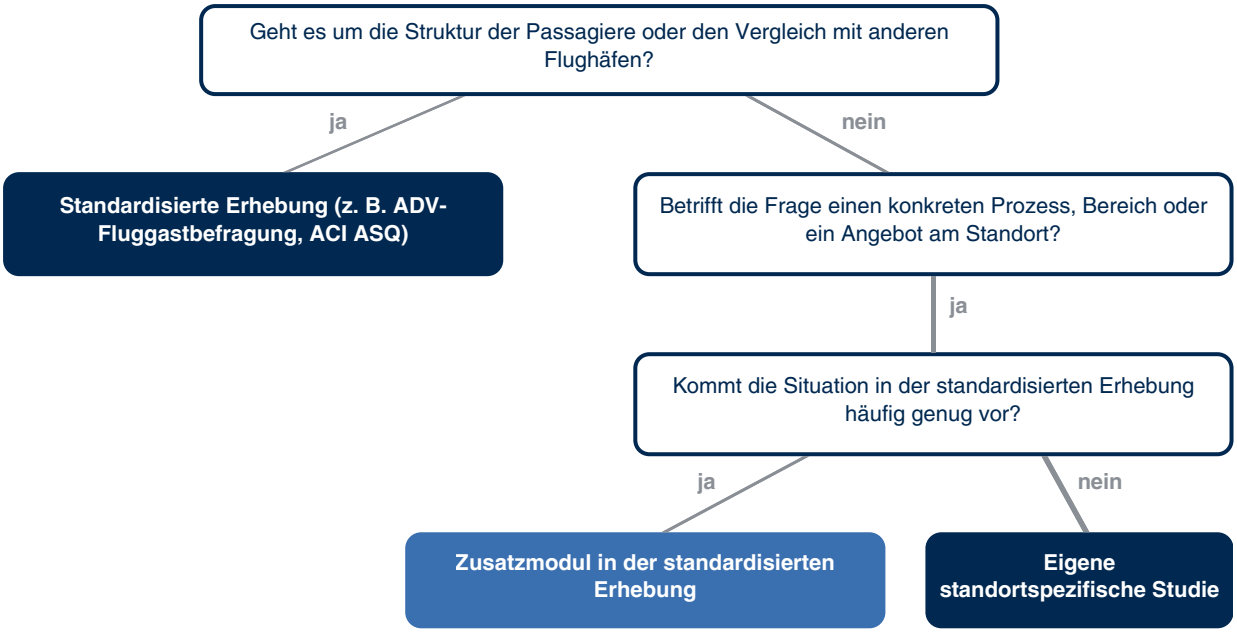

Abb.: Entscheidungshilfe zur Wahl des Studientyps

Nicht Gegenstand dieses Whitepapers

Befragungen sind nicht das einzige Verfahren, mit dem sich das Verhalten und Erleben von Passagieren untersuchen lässt. Eye-Tracking zeigt, welche Wegweiser und Anzeigen tatsächlich wahrgenommen werden, Frequenzmessungen und Beobachtungen zeigen, wie sich Passagierströme durch das Terminal bewegen, und begleitete Check-ins, vergleichbar mit begleiteten Einkäufen im Handel, zeigen im direkten Miterleben, wo Passagiere zögern oder Hilfe brauchen. Diese Verfahren liefern wertvolle Ergänzungen und werden in diesem Whitepaper dort erwähnt, wo sie eine Befragung sinnvoll ergänzen. Im Mittelpunkt steht jedoch die Befragung. Eine ausführliche Darstellung des Eye-Trackings bietet das IWD-Whitepaper „Eye-Tracking im Handel“.

4. Methodische Grundlagen

Die Qualität einer Fluggastbefragung entsteht in vier Planungsentscheidungen: wie befragt wird, wer befragt wird, wie viele Interviews nötig sind und was gefragt wird. Hinzu kommt die Umsetzung im Feld, die am Flughafen mehr Einfluss auf die Datenqualität hat als in den meisten anderen Befragungssituationen.

4.1 Erhebungsformen: CAPI, CASI, Mixed-Mode und Online

Am Flughafen kommen im Kern drei Erhebungsformen zum Einsatz. Beim **CAPI** (Computer Assisted Personal Interviewing) führt ein Interviewer das Gespräch und erfasst die Antworten auf dem Tablet. Beim **CASI** (Computer Assisted Self Interviewing) füllt der Passagier den Fragebogen selbst auf einem Tablet aus, das ihm ein Interviewer übergibt. **Online- und QR-Code-Befragungen** laden Passagiere über Aufsteller, Anzeigen oder digitale Kanäle zur Teilnahme ein, ohne dass eine persönliche Ansprache stattfindet.

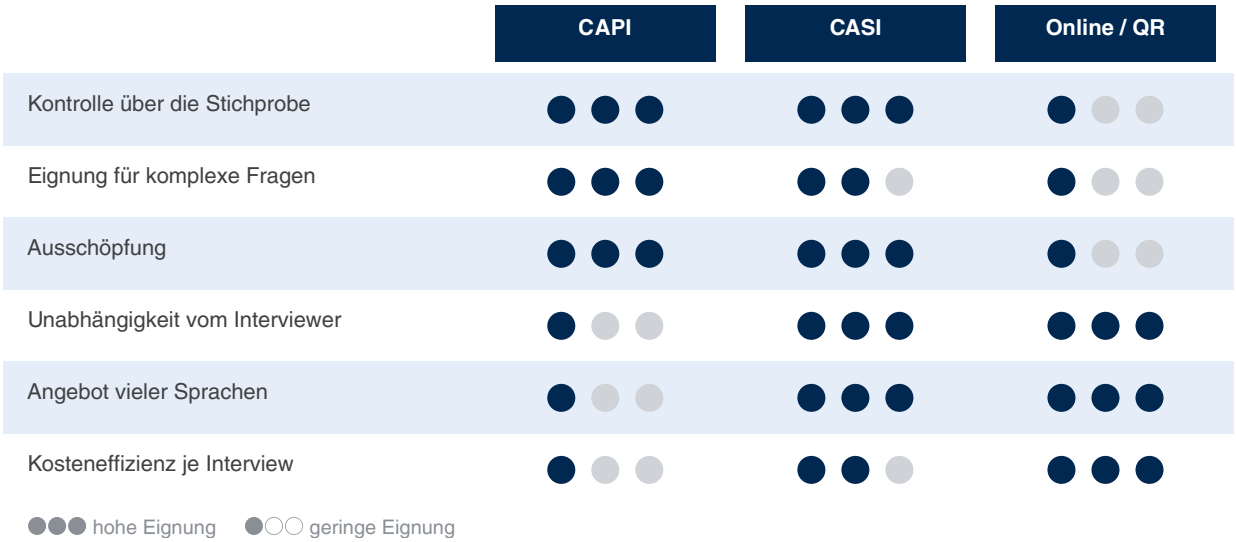

Abb.: Stärken der Erhebungsformen im Vergleich

CAPI bietet die größte Kontrolle. Der Interviewer setzt die Auswahlregel um, erklärt Fragen, die ohne Erläuterung zu Fehlern führen würden, und stellt sicher, dass das Interview vollständig ist. Das ist unverzichtbar für Strukturfragen, bei denen Präzision zählt, etwa ob ein Passagier nur umsteigt oder seine Reise an diesem Flughafen beginnt und ob die angegebene Postleitzahl zum Wohnort oder zum Hotel gehört. Der Nachteil ist neben dem höheren Aufwand ein systematischer Einfluss auf die Antworten, denn in Anwesenheit eines Interviewers fallen Bewertungen tendenziell freundlicher aus. Eine umfangreiche Mixed-Mode-Studie zeigt, dass Befragte, denen die Fragen vorgelesen werden, deutlich häufiger sehr positive Zufriedenheitsurteile abgeben als beim Selbstausfüllen (Dillman et al. 2009).

CASI verlagert die Beantwortung auf den Passagier. Das reduziert den Interviewereinfluss gerade bei Zufriedenheits- und Bewertungsfragen, erlaubt längere Fragenblöcke und mehrere Sprachversionen parallel. Ein Interviewer kann zudem mehrere Passagiere gleichzeitig betreuen. Die Limitation liegt in der Verständlichkeit, denn was im Selbstausfüller missverstanden wird, fällt niemandem auf. CASI setzt deshalb einen besonders sorgfältig getesteten Fragebogen voraus.

Online- und QR-Code-Befragungen sind am günstigsten und liefern kontinuierlich Rückmeldungen, auch außerhalb klassischer Feldzeiten. Sie geben aber die Kontrolle darüber ab, wer teilnimmt. Erfahrungsgemäß antworten vor allem Passagiere mit besonders positiven oder besonders negativen Erlebnissen. Für repräsentative Aussagen über die Passagierstruktur oder für belastbare Zufriedenheitswerte sind sie deshalb nicht geeignet. Ihre Stärke liegt als Frühwarnsystem. Sie zeigen schnell, wenn an einem Kontaktpunkt zwischen Passagier und Flughafen, einem sogenannten Touchpoint wie Check-in oder Sicherheitskontrolle, etwas aus dem Ruder läuft.

Mixed-Mode: die Kombination als Standard

In der Praxis hat sich die Kombination von CAPI und CASI bewährt. Der Interviewer spricht den Passagier nach Auswahlregel an, stellt im CAPI-Teil die Einstiegs- und Strukturfragen und übergibt dann das Tablet für den Bewertungsteil. So verbinden sich die Stärken beider Verfahren, also kontrollierte Stichprobe und präzise Strukturdaten auf der einen Seite, geringerer Interviewereinfluss und Sprachvielfalt auf der anderen.

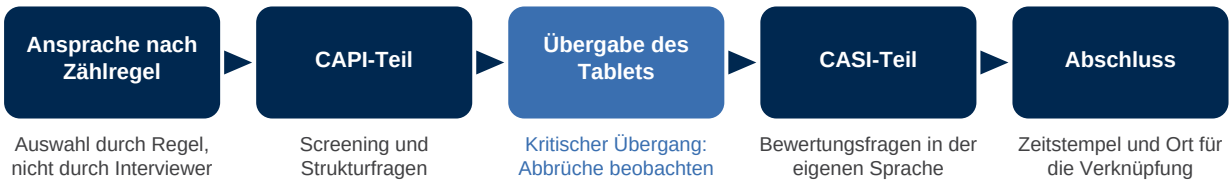

Abb.: Ablauf eines Mixed-Mode-Interviews

Bei standardisierten Erhebungen trägt der CAPI-Teil die Strukturfragen, deren Qualität über die spätere Gewichtung entscheidet. Bei standortspezifischen Studien übernimmt er vor allem das Screening, also kurze Einstiegsfragen, etwa ob die Person den untersuchten Bereich tatsächlich genutzt hat, während der Schwerpunkt im CASI-Teil liegt.

Moduseffekte: Der Erhebungsweg beeinflusst das Ergebnis

Die Kombination hat eine Konsequenz, die häufig unterschätzt wird. Bewertungen aus dem CAPI-Teil sind nicht ohne Weiteres mit Bewertungen aus dem CASI-Teil oder aus einer Online-Befragung vergleichbar. Wer die Zuordnung von Fragen zu Erhebungswegen verändert, verändert damit auch die Ergebnisse. Daraus folgen zwei Regeln. Erstens sollte festgelegt werden, welche Fragen in welchem Modus gestellt werden, und diese Zuordnung sollte über alle Erhebungswellen konstant bleiben. Zweitens sollte ein Moduswechsel in einer laufenden Zeitreihe nie ohne Parallelerhebung erfolgen, die den Effekt des Wechsels messbar macht.

USE CASE: Die Zufriedenheit sinkt, aber am Flughafen hat sich nichts verändert

Ausgangslage: Ein Flughafen stellte seine laufende Passagierbefragung von einem vollständig interviewergeführten Verfahren auf Mixed-Mode um, um mehr Sprachen anbieten zu können. In der ersten Welle nach der Umstellung lagen fast alle Zufriedenheitswerte spürbar unter dem Vorjahresniveau, obwohl es keine operativen Veränderungen gegeben hatte.

Methodisches Vorgehen: In der folgenden Welle wurde ein Teil der Stichprobe weiterhin im alten Verfahren befragt, der andere im neuen. Beide Teilstichproben waren nach derselben Auswahlregel gezogen und über dieselben Zeitfenster verteilt.

Ergebnis: Die Differenz zwischen den Teilstichproben entsprach nahezu vollständig dem scheinbaren Rückgang. Der Einbruch war ein Moduseffekt, kein verändertes Erleben.

Konsequenz: Die Zeitreihe wurde mit einem dokumentierten Bruch fortgeführt, die Vorjahreswerte für den Vergleich entsprechend eingeordnet. Ohne die Parallelerhebung hätte der Flughafen Maßnahmen gegen ein Problem ergriffen, das es nicht gab.

4.2 Stichprobe, Quotierung und Auswahl der Passagiere

Die Stichprobe ist der Punkt, an dem sich die beiden Studientypen am deutlichsten unterscheiden. Die folgende Übersicht stellt die Planungsschritte gegenüber.

Planungsschritt	Standardisierte Erhebung	Standortspezifische Studie
Grundgesamtheit	Alle abfliegenden Passagiere eines Zeitraums	Passagiere, die den untersuchten Prozess, Bereich oder Service nutzen oder nutzen könnten
Erhebungsort	Wartebereiche im Abflugbereich, verteilt über alle Terminals und Pieranlagen	Am untersuchten Touchpoint oder dort, wo betroffene Passagiere erreichbar sind
Zeitplanung	Über Flugplanperioden, Wochentage und Tageszeiten verteilt, einschließlich der Randzeiten am frühen Morgen und späten Abend	Konzentriert auf die relevanten Situationen, z. B. Spitzenzeiten oder Zeiträume vor und nach einer Veränderung
Quotierung	SOLL-Verteilung nach Zielflughafen, Airline und Zeitfenster auf Basis von PAX-Zahlen	Nach Merkmalen, die für die Fragestellung relevant sind, z. B. Nutzer und Nichtnutzer, Auslastung, Reiseanlass
Auswahl im Feld	Zufallsgesteuert nach fester Zählregel, eine Person je Reisegruppe	Zufallsgesteuert nach fester Zählregel, ergänzt um ein Screening
Gewichtung	Anpassung an das reale Passagieraufkommen	In der Regel keine Hochrechnung; Anpassung nur bei bewusst disproportionaler Erhebung

Standardisierte Erhebungen: Quoten nach Verkehrsstruktur

Ziel einer standardisierten Erhebung ist ein Abbild der gesamten Passagierstruktur. Grundlage der Planung ist deshalb eine SOLL-Verteilung der Interviews, die aus dem Flugplan und den Passagierzahlen vergleichbarer Vorperioden abgeleitet wird. Sie legt fest, welcher Anteil der Interviews auf welchen Zielflughafen, welche Airline, welchen Wochentag und welches Zeitfenster entfällt. Während der Feldphase wird die erreichte IST-Verteilung laufend mit dieser SOLL-Verteilung abgeglichen, sodass Abweichungen früh korrigiert werden können, etwa wenn Randzeiten unterbesetzt sind.

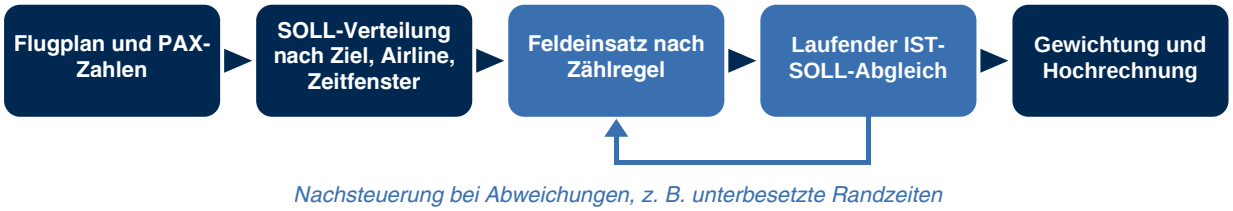

Abb.: Steuerung der Stichprobe bei standardisierten Erhebungen

Methodisch handelt es sich um eine sogenannte Klumpenstichprobe. Befragt wird nicht aus einer Liste aller Passagiere, sondern an einzelnen Abflügen. Passagiere desselben Fluges ähneln sich aber in Reiseziel, Reiseanlass und oft auch Herkunft. Viele Interviews an wenigen Flügen liefern deshalb weniger Information als dieselbe Zahl von Interviews, verteilt über viele Flüge. Eine gute Einsatzplanung begrenzt die Zahl der Interviews je Abflug und verteilt die Stichprobe breit über den Flugplan.

Werden die Interviews genau nach den Passagieranteilen verteilt, entfallen auf kleine Zielflughäfen oder Airlines allerdings nur wenige Interviews. Das ist kein Problem für Aussagen über den Gesamtflughafen, wohl aber für Aussagen über genau diese Verbindungen, die strategisch oft besonders interessant sind, etwa neue Strecken oder Märkte, in denen der Flughafen wachsen will. Diese Verbindungen werden deshalb gezielt überproportional befragt und in der Auswertung auf ihren tatsächlichen Anteil zurückgewichtet.

Rechenbeispiel: Warum kleine Strecken aufgestockt werden

Ein Flughafen führt pro Jahr 2.000 Interviews durch. Eine neu aufgenommene Verbindung hat einen Anteil von 2 % am Passagieraufkommen. Bei proportionaler Verteilung entfallen darauf rund 40 Interviews im Jahr, also etwa zehn pro Quartal. Ein Anteilswert aus 40 Interviews, etwa der Anteil der Geschäftsreisenden, kann mit 95-prozentiger Sicherheit um bis zu rund ±15 Prozentpunkte vom tatsächlichen Wert abweichen, bei zehn Interviews um mehr als ±30 Prozentpunkte. Aussagen über Reiseanlass oder Einzugsgebiet dieser Passagiere sind damit kaum möglich. Wird die Verbindung auf 150 Interviews im Jahr aufgestockt, sinkt die Unsicherheit auf rund ±8 Prozentpunkte. In der Gesamtauswertung erhalten diese Interviews einen entsprechend geringeren Gewichtungsfaktor, sodass das Gesamtergebnis unverzerrt bleibt.

Standortspezifische Studien: Stichprobe aus der Fragestellung

Standortspezifische Studien leiten ihre Stichprobe nicht aus der Verkehrsstruktur ab, sondern aus der Fragestellung. Grundgesamtheit sind diejenigen Passagiere, die den untersuchten Prozess, Bereich oder Service durchlaufen haben oder hätten nutzen können. Befragt wird dort, wo diese Personen erreichbar sind, etwa unmittelbar nach der Sicherheitskontrolle, im Wartebereich eines bestimmten Piers, in der Ankunftshalle oder bei Passagieren eines Fluges, dessen Gate kurzfristig gewechselt wurde.

Gesteuert wird nach den Situationen, die für die Fragestellung relevant sind. Wer wissen will, wie Wartezeit bei hoher Auslastung erlebt wird, muss gezielt in Spitzenzeiten erheben (siehe Use Case in Abschnitt 6.2). Bei Fragestellungen zu Angeboten gehören Nichtnutzer in die Stichprobe. Wer nur die Gäste eines Restaurants befragt, erfährt, was die ohnehin Überzeugten schätzen, aber nicht, was alle anderen abhält. Eine Hochrechnung auf die gesamte Passagierstruktur ist bei solchen Studien in der Regel weder nötig noch sinnvoll.

Auswahl der Passagiere im Feld

Unabhängig vom Studientyp gilt eine Regel, die über die Qualität der gesamten Studie entscheidet. Nicht der Interviewer wählt aus, wen er anspricht, sondern ein festes Verfahren. Üblich ist eine Zählregel, nach der etwa jede n-te Person angesprochen wird, die einen definierten Punkt passiert oder sich im Wartebereich setzt. Bei Reisegruppen wird nur eine Person befragt, deren Auswahl ebenfalls einer Regel folgt, zum Beispiel die Person, die als Nächstes Geburtstag hat.

Denn ohne Regel sprechen Interviewer unwillkürlich Menschen an, die zugänglich, entspannt und gesprächsbereit wirken. Gestresste Geschäftsreisende, Familien mit müden Kindern oder Passagiere, die nicht Deutsch sprechen, wären dann systematisch unterrepräsentiert, und gerade ihre Erfahrungen sind oft besonders aufschlussreich. Bei standortspezifischen Studien folgt auf die Ansprache häufig ein kurzes Screening. Wichtig ist, dass es nach der Ansprache gemäß Zählregel erfolgt und nicht durch die Einschätzung des Interviewers ersetzt wird. Wer nur Passagiere mit Einkaufstüten anspricht, befragt Käufer, nicht Besucher.

4.3 Fallzahl und Genauigkeit

Die Frage nach der richtigen Anzahl der Interviews gehört zu den häufigsten in der Planung, und sie wird oft zu früh gestellt. Denn die entscheidende Frage ist nicht, wie viele Interviews insgesamt geführt werden, sondern wie viele in der kleinsten Gruppe, über die eine belastbare Aussage getroffen werden soll. Eine große Fallzahl schützt nicht vor einer schlechten Stichprobe. 2.000 Interviews können für den Gesamtflughafen ausreichen und für die wichtigste Teilgruppe dennoch kaum belastbar sein.

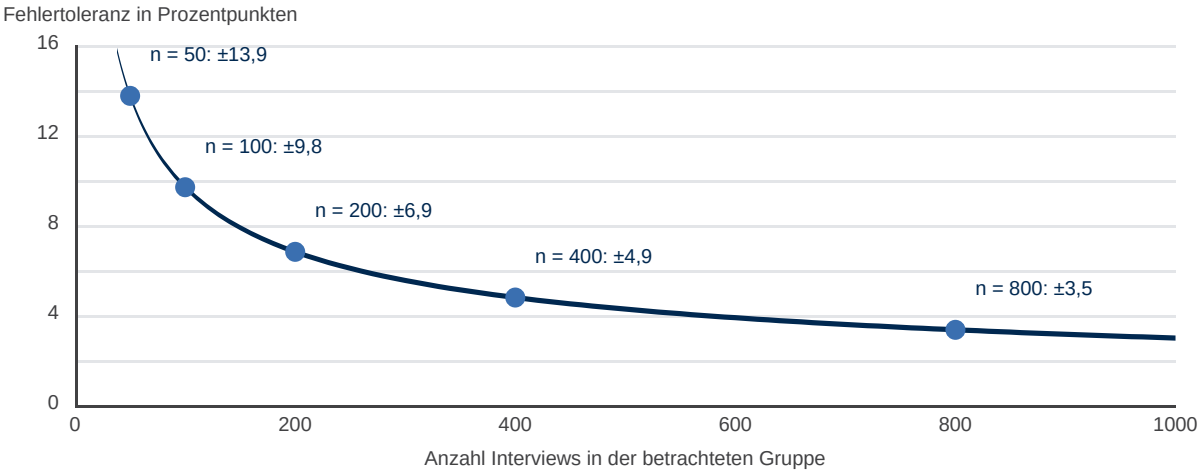

Abb.: Fehlertoleranz eines Anteilswerts nach Fallzahl (Konfidenzniveau 95 %, Anteilswert 50 %, einfache Zufallsstichprobe)

Die Grafik zeigt einen für die Planung zentralen Zusammenhang. Die Genauigkeit wächst nicht proportional zur Fallzahl, sondern deutlich langsamer. Um die Fehlertoleranz zu halbieren, braucht man viermal so viele Interviews. Der Schritt von 50 auf 200 Interviews halbiert die Fehlertoleranz, der Schritt von 400 auf 800 Interviews verbessert sie nur noch um knapp eineinhalb Prozentpunkte. Geben zum Beispiel bei 200 Interviews 50 % der Befragten an, privat zu reisen, liegt der tatsächliche Anteil unter allen Passagieren mit 95-prozentiger Sicherheit zwischen etwa 43 und 57 %. Für Vergleiche zwischen Gruppen gilt zusätzlich, dass ein Unterschied von fünf Prozentpunkten zwischen zwei Gruppen mit je 150 Interviews noch im Bereich des Zufalls liegt. Wer solche Unterschiede nachweisen will, braucht deutlich größere Gruppen.

Bei standardisierten Erhebungen kommt ein weiterer Faktor hinzu. Gewichtung und die Befragung an einzelnen Flügen verringern die sogenannte effektive Fallzahl, also die Zahl der Interviews, die statistisch tatsächlich zählt. Stark unterschiedliche Gewichtungsfaktoren, etwa durch umfangreiche Aufstockungen, lassen 2.000 Interviews statistisch wie deutlich weniger wirken. Dieser sogenannte Designeffekt (Kish 1965) sollte in der Planung berücksichtigt und in der Ergebnisdarstellung transparent gemacht werden.

Studienziel	Standortspezifische Studie	Standardisierte Erhebung
Muster erkennen, Schwachstellen identifizieren	mind. 100 bis 150 Interviews je untersuchtem Bereich	ergibt sich aus der Quotenplanung
Unterschiede zwischen Gruppen oder Wellen nachweisen	mind. 300 bis 400 Interviews je Vergleichsgruppe	aufgestockte Teilgruppen mit ausreichender Fallzahl je Auswertungsperiode

Tab.: Orientierungswerte für die Planung. Die tatsächlich benötigte Fallzahl hängt von der erwarteten Effektgröße und der Zahl der Vergleichsgruppen ab.

4.4 Fragebogen: standardisierter Kern, gezielte Module

Ein guter Flughafenfragebogen ist kurz, aber nicht dünn. Er muss an einem Ort funktionieren, an dem die Befragten jederzeit aufgerufen werden können, und dennoch genug Tiefe liefern, um Ursachen zu erkennen. Bewährt hat sich ein modularer Aufbau aus einem stabilen Kern und wechselnden Vertiefungsmodulen.

Der Kern umfasst die Strukturmerkmale, die für jede Segmentierung gebraucht werden (Reiseanlass, Reisehäufigkeit, Umsteigesituation, Herkunft, Begleitung), die Bewertung der zentralen Touchpoints und die Gesamturteile. Er bleibt über alle Wellen unverändert, denn jede Änderung im Wortlaut, in der Skala oder in der Reihenfolge kann die Zeitreihe brechen. Module vertiefen wechselnde Themen, etwa einen neuen Gastronomiebereich, eine umgebaute Wegeführung oder die Nutzung digitaler Services. Standortspezifische Studien sollten die zentralen Strukturmerkmale in identischer Formulierung wie die standardisierte Erhebung erfassen. So lassen sich ihre Ergebnisse denselben Zielgruppen zuordnen.

Länge und Ort gehören zusammen.

Im Wartebereich vor dem Boarding ist ein Interview von etwa zehn bis fünfzehn Minuten realistisch. Direkt nach der Sicherheitskontrolle oder in der Ankunftshalle sollte ein Fragebogen deutlich kürzer sein, weil Passagiere dort weitergehen wollen. Jede zusätzliche Frage erhöht zudem das Risiko von Abbrüchen, und diese treffen nicht alle Passagiere gleich (siehe Abschnitt 6.1).

Erlebnisse statt Pauschalurteile.

Allgemeine Zufriedenheitsfragen liefern Durchschnittswerte, konkrete Erlebnisfragen liefern Ansatzpunkte. Die Frage, wie lange jemand ungefähr an der Sicherheitskontrolle gewartet hat, ist aufschlussreicher als die Frage, wie zufrieden er mit der Wartezeit war, weil sie sich mit der gemessenen Wartezeit desselben Zeitfensters vergleichen lässt. Ebenso zeigt die Frage, ob jemand an einer bestimmten Stelle unsicher war, welchen Weg er nehmen sollte, mehr als eine Note für die Beschilderung.

Einheitliche Skalen.

Alle Bewertungsfragen des Kerns sollten dieselbe Skala verwenden, damit Touchpoints untereinander vergleichbar sind. Die Weiterempfehlung wird als Net Promoter Score (Reichheld 2003) auf der dafür vorgesehenen Skala von 0 bis 10 erhoben. Für die Berichterstattung lohnt es sich, neben Mittelwerten auch Top-Box-Anteile auszuweisen, also den Anteil der sehr zufriedenen Passagiere. Sie reagieren oft empfindlicher auf Veränderungen als der Mittelwert, weil sich Verbesserungen häufig zuerst darin zeigen, dass aus zufriedenen Passagieren begeisterte werden.

Pretest unter realen Bedingungen.

Jeder neue Fragebogen und jede neue Sprachversion sollte vor dem Feldstart am Flughafen selbst getestet werden, mit echten Passagieren, im echten Zeitfenster. Nur so zeigt sich, ob die Interviewlänge passt, ob Begriffe wie Transfer, Boarding oder Sicherheitskontrolle in allen Sprachen verstanden werden und ob die Fragenführung funktioniert, wenn Passagiere den Fragebogen selbst ausfüllen.

4.5 Feldarbeit und Qualitätssicherung

Am Flughafen entscheidet die Feldarbeit stärker über die Datenqualität als in den meisten anderen Befragungssituationen. Die beste Stichprobenplanung nützt wenig, wenn Interviewer die Zählregel, also das feste Verfahren, nach dem zum Beispiel jede n-te Person angesprochen wird (siehe Abschnitt 4.2), unter Zeitdruck nicht einhalten, Randzeiten unbesetzt bleiben oder Gate-Wechsel die Einsatzplanung durcheinanderbringen.

Einsatzplanung entlang des Flugplans.

Schichten und Einsatzorte werden aus dem Flugplan abgeleitet, damit die Interviewer dort sind, wo die geplanten Abflüge stattfinden. Da sich der Flugplan täglich verändert, braucht es eine Feldleitung vor Ort, die Verspätungen, Gate-Wechsel und Annullierungen verfolgt und die Teams kurzfristig umsetzt.

Akkreditierung und Sicherheitsauflagen.

Flughafenausweise, Zuverlässigkeitsüberprüfungen, Einweisungen in Sicherheitsregeln und die Abstimmung mit Terminalbetrieb und Bundespolizei müssen in den Projektplan eingerechnet werden. Wechsel im Interviewerteam sind wegen dieses Vorlaufs nur eingeschränkt möglich. Ein stabiler Stamm erfahrener Interviewer ist deshalb ein Qualitätsfaktor.

Anforderungen an Interviewer.

Neben Sprachkenntnissen brauchen Interviewer am Flughafen ein gutes Gespür für Situationen. Sie müssen erkennen, wann ein Passagier Zeit hat und wann nicht, ein Interview beim Boardingaufruf souverän beenden und die Zählregel auch dann einhalten, wenn die nächste Person wenig gesprächsbereit wirkt.

Laufende Kontrollen.

Tabletgestützte Erhebungen erlauben Kontrollen während der Feldphase, nicht erst danach. Zeitstempel und Interviewdauer zeigen auffällig schnell durchlaufene Fragebögen, Abbruchquoten je Interviewer, Zeitfenster und Fragebogenteil zeigen, wo der Ablauf nicht funktioniert. Ungewöhnlich homogene Antwortmuster oder eine deutlich abweichende Verweigerungsquote einzelner Interviewer werden unmittelbar geprüft. Im Mixed-Mode ist die Abschlussquote am Übergang vom CAPI- zum CASI-Teil ein besonders aussagekräftiger Indikator.

Datenschutz.

Passagiere werden vor Beginn über Zweck, Freiwilligkeit und Anonymität informiert. Angaben, die Rückschlüsse auf eine Person zulassen könnten, etwa Flugnummer und Postleitzahl in Kombination, werden nur so detailliert erhoben, wie es die Auswertung erfordert, und gemäß DSGVO verarbeitet. Tablets speichern verschlüsselt und übertragen regelmäßig, damit bei Verlust oder Ausfall keine Daten verloren gehen oder in falsche Hände geraten.

Wichtig für die Planung

Für Zugangsberechtigungen und Zuverlässigkeitsüberprüfungen sollte ein Vorlauf von mehreren Wochen eingeplant werden. Nicht jedes angesprochene Interview wird abgeschlossen, denn Boardingaufrufe, Gate-Wechsel und Zeitdruck führen zu Abbrüchen, die in der Einsatzplanung einkalkuliert werden müssen. Randzeiten am frühen Morgen und späten Abend sind in der Feldplanung am schwersten zu besetzen und fallen deshalb am häufigsten aus der Stichprobe heraus. Sie sollten von Beginn an fest eingeplant werden.

5. Kennzahlen und Auswertung

5.1 Kennzahlen-Systematik: Struktur, Erleben, Wirkung

Aus einer Fluggastbefragung lassen sich zahlreiche Kennzahlen ableiten. Damit sie nicht als unverbundene Liste nebeneinanderstehen, hilft eine Ordnung nach drei Fragen. Wer nutzt den Flughafen, wie werden seine Prozesse und Angebote erlebt, und was folgt daraus für das Verhalten der Passagiere? Erst das Zusammenspiel der drei Ebenen macht Ergebnisse erklärbar. Strukturdaten zeigen, für wen ein Befund gilt, Erlebenskennzahlen zeigen, wo die Ursachen liegen, Wirkungskennzahlen zeigen, was auf dem Spiel steht.

Struktur *Wer nutzt den Flughafen?*	**Erleben** *Wie werden Prozesse wahrgenommen?*	**Wirkung** *Was folgt daraus?*
Einzugsgebiet *Herkunft der Passagiere*	**Zufriedenheit je Touchpoint** *Bewertung einzelner Schritte*	**Gesamtzufriedenheit** *Übergreifendes Urteil*
Reiseanlass und Reiseziel *Warum und wohin gereist wird*	**Erlebte Wartezeit** *Geschätzte Minuten*	**Weiterempfehlung (NPS)** *Verhaltensnahe Größe*
Anreise und Verkehrsmittel *Wie Passagiere ankommen*	**Wahrnehmungsdifferenz** *Erlebt minus gemessen*	**Nutzungsquote Non-Aviation** *Anteil der Nutzer*
Umsteigesituation *Umstieg oder Reisebeginn hier*	**Orientierungssicherheit** *Unsicherheit an Entscheidungspunkten*	**Nichtnutzungsgründe** *Barrieren der Nichtnutzer*
Zeitpuffer vor Abflug *Verfügbare Zeit im Terminal*	**Stressempfinden** *Subjektive Belastung*	**Flughafenwahl** *Gründe und Alternativen*

Abb.: Kennzahlen einer Fluggastbefragung in drei Kategorien

Strukturkennzahlen beschreiben, wer den Flughafen nutzt: Einzugsgebiet, Reiseanlass und Reiseziel, Anreiseverkehrsmittel, Umsteigesituation und den Zeitpuffer vor dem Abflug. Sie sind die Grundlage für Marketing, Verkehrsanbindung und Parkraumplanung und für die Argumentation gegenüber Airlines. Der Zeitpuffer zeigt zudem, wie viel Zeit Passagiere nach der Sicherheitskontrolle haben und damit, wie groß das Non-Aviation-Potenzial überhaupt ist.

Erlebenskennzahlen zeigen, wie Prozesse und Angebote wahrgenommen werden: Zufriedenheit je Touchpoint, erlebte Wartezeit, Orientierungssicherheit an konkreten Entscheidungspunkten und Stressempfinden. Sie lokalisieren Schwachstellen und zeigen Handlungsbedarf jenseits der Durchlaufzeit, etwa bei Information, Warteschlangenführung oder Beschilderung.

Wirkungskennzahlen zeigen, was aus dem Erleben folgt: Gesamtzufriedenheit und Weiterempfehlung (NPS), die Nutzung von Retail und Gastronomie, Gründe für die Nichtnutzung und die Gründe, aus denen Passagiere diesen Flughafen gewählt haben. Der NPS wird dabei nach Reichheld (2003) als verhaltensnähere Größe als die Zufriedenheit verstanden, die Flughafenwahl als Indikator für die Wettbewerbsposition im Einzugsgebiet.

Eine Kennzahl verdient besondere Aufmerksamkeit, weil sie in klassischen Kennzahlensystemen fehlt, nämlich die Wahrnehmungsdifferenz. Sie entsteht, wenn die erfragte Wartezeit mit der gemessenen Durchlaufzeit desselben Ortes und Zeitfensters verglichen wird. Liegt die erlebte Wartezeit deutlich über der gemessenen, deutet das auf fehlende Information, unübersichtliche Wartebereiche oder Zeitdruck hin. Diese Ursachen lassen sich oft mit geringem Aufwand beheben, während eine Verkürzung der tatsächlichen Wartezeit zusätzliches Personal oder zusätzliche Kontrollspuren erfordert.

5.2 Von Daten zu Erkenntnissen

Zwischen Erhebung und Entscheidung liegen mehrere Analyseschritte, die aus Durchschnittswerten eine Entscheidungsgrundlage machen.

Segmentierung

Durchschnittswerte über alle Passagiere verdecken oft das Wesentliche. Ein Flughafen mit guter Gesamtzufriedenheit kann bei Familien mit Kindern, bei Umsteigern oder bei Passagieren in der Morgenspitze deutliche Schwächen haben. Die Segmentierung nach Strukturmerkmalen ist deshalb der erste Auswertungsschritt. Noch aussagekräftiger als einzelne Merkmale sind Typologien, die mehrere Merkmale verbinden. So umfasst die Kategorie Geschäftsreisender den Tagespendler mit minimalem Zeitpuffer ebenso wie den Langstreckenreisenden mit mehrstündigem Aufenthalt, deren Erwartungen an Abläufe, Lounges und Angebote kaum etwas gemeinsam haben.

Verknüpfung mit operativen Kennzahlen

Den größten Erkenntniswert entfalten Befragungsdaten, wenn sie mit operativen Kennzahlen desselben Ortes und Zeitraums verknüpft werden. Voraussetzung ist, dass jedes Interview mit Zeitstempel und Befragungsort erfasst wird. Dann lässt sich jede Antwort den Durchlaufzeiten, Auslastungen oder Verspätungen des entsprechenden Zeitfensters zuordnen. So wird sichtbar, ab welcher Wartezeit die Zufriedenheit tatsächlich einbricht, ob eine Verspätung das Urteil über den gesamten Flughafen beeinflusst oder ob Umsätze in einem Bereich mit dem erlebten Zeitdruck zusammenhängen.

Treiberanalyse: Welche Touchpoints das Gesamturteil tragen

Die naheliegende Auswertung einer Befragung ist eine Rangliste, die zeigt, welcher Touchpoint am schlechtesten bewertet wird. Für Entscheidungen ist diese Frage jedoch unvollständig. Entscheidend ist, welche Touchpoints das Gesamturteil oder die Weiterempfehlung tatsächlich beeinflussen. Genau das beantwortet die Treiberanalyse. Sie schätzt statistisch, wie stark die Bewertung einzelner Prozessschritte mit der Gesamtzufriedenheit oder dem NPS zusammenhängt. Weil die Bewertungen einzelner Touchpoints untereinander stark korrelieren, empfehlen sich Verfahren, die den Beitrag jedes Treibers auch bei solchen Überschneidungen stabil schätzen, etwa das Verfahren der relativen Gewichte (Johnson 2000).

Stellt man den Einfluss eines Touchpoints seiner aktuellen Bewertung gegenüber, entsteht eine Prioritätenmatrix. Sie ordnet jeden Touchpoint einem von vier Feldern zu und zeigt, wo Investitionen die größte Wirkung entfalten.

Abb.: Prioritätenmatrix aus Treiberanalyse und Bewertung (schematische Darstellung)

USE CASE: Gute Gesamtnote, schwache Weiterempfehlung

Ausgangslage: An einem Flughafen lag die Gesamtzufriedenheit über mehrere Erhebungswellen stabil im guten Bereich. Der Net Promoter Score entwickelte sich dagegen deutlich schwächer, als die Gesamtnote erwarten ließ.

Methodisches Vorgehen: Für Gesamtzufriedenheit und NPS wurde jeweils eine eigene Treiberanalyse über alle bewerteten Touchpoints gerechnet.

Ergebnis: Der Check-in wurde von den Passagieren, die ihn am Schalter genutzt hatten, konstant unterdurchschnittlich bewertet, beeinflusste die Gesamtzufriedenheit aber kaum, weil er nur einer von vielen Prozessschritten war. Für die Weiterempfehlung war er dagegen einer der stärksten Treiber. In einer einfachen Rangliste wäre er als einer von mehreren mittelmäßigen Punkten erschienen, in der Prioritätenmatrix für den NPS lag er klar im Feld der Handlungsfelder. Die offenen Nennungen erklärten das, denn der Check-in prägt den ersten Eindruck der Reise, und ein schlechter Start färbt das Urteil darüber, ob man den Flughafen anderen empfehlen würde. Das deckt sich mit der Beobachtung von Maister (1985), dass die frühen Phasen einer Dienstleistung das Gesamterleben überproportional prägen.

Konsequenz: Der Flughafen priorisierte Maßnahmen am Check-in vor Verbesserungen an Touchpoints, die in der Rangliste schlechter abgeschnitten hatten, aber für die Weiterempfehlung kaum eine Rolle spielten.

Benchmarking und Zeitreihen

Einzelne Werte gewinnen ihre Bedeutung erst im Vergleich. Standardisierte Programme wie ACI ASQ ermöglichen die Einordnung der eigenen Servicequalität im internationalen Vergleich, die ADV-Fluggastbefragung den Vergleich der Passagierstruktur mit anderen Flughäfen. Mit Benchmarks vergleichbar sind Werte allerdings nur, wenn Frageformulierung, Skala und Erhebungsweg übereinstimmen. Mindestens ebenso wertvoll ist die eigene Zeitreihe, sofern Fragebogenkern und Erhebungsweg stabil gehalten werden.

Gewichtung und Hochrechnung bei standardisierten Erhebungen

Bei standardisierten Erhebungen wird die Stichprobe in der Auswertung an die tatsächliche Verkehrsstruktur angepasst. Aufgestockte Verbindungen erhalten einen geringeren, unterbesetzte Zeitfenster einen höheren Gewichtungsfaktor. Über einen Hochrechnungsfaktor lassen sich die Ergebnisse zudem auf das reale Passagieraufkommen beziehen, etwa um die Zahl der Passagiere aus einer bestimmten Region zu schätzen. Gewichtung ist jedoch kein Ersatz für eine gute Feldarbeit. Sie kann eine dünn besetzte Gruppe rechnerisch vergrößern, aber keine fehlende Information ergänzen. Zehn Interviews bleiben zehn Interviews, auch wenn sie für 50.000 Passagiere stehen.

6. Methodische Herausforderungen und Limitationen

Jede Erhebungsmethode unterliegt Limitationen. Am Flughafen sind einige davon besonders ausgeprägt. Wer sie kennt, kann sie in der Planung berücksichtigen und in der Interpretation richtig einordnen. Während Kapitel 4 beschreibt, wie eine Fluggastbefragung aufgebaut wird, zeigt dieses Kapitel, wo die typischen Fallstricke liegen und wie man ihnen begegnet.

6.1 Zeitdruck und Befragungssituation

Die Befragungssituation am Flughafen ist nie neutral. Passagiere sind auf dem Weg zu einem Flug, und ihre Aufmerksamkeit ist geteilt zwischen dem Interview, der Anzeigetafel und der Sorge, den Aufruf zu verpassen. Das hat drei Folgen für die Daten.

Erstens beeinflusst der Befragungszeitpunkt das Urteil. Wer unmittelbar nach einer langen Wartezeit befragt wird, bewertet anders als jemand, der die Kontrolle vor einer Stunde passiert hat und inzwischen entspannt im Café sitzt. Beides ist legitim, aber es misst Unterschiedliches, nämlich das unmittelbare Erleben oder das rückblickende Urteil. Die Wahl des Befragungsortes muss deshalb zur Fragestellung passen und über alle Wellen konstant bleiben.

Zweitens betreffen Abbrüche nicht alle Passagiere gleich. Wer spät am Gate ankommt, hat weniger Zeit und bricht häufiger ab. Das sind oft genau die Passagiere, die unter Zeitdruck standen und die Prozesse als besonders belastend erlebt haben. Ein langer Fragebogen verzerrt die Stichprobe deshalb in Richtung der entspannten Passagiere und damit in Richtung zu positiver Ergebnisse. Kurze Fragebögen und die Auswertung von Abbrüchen nach Ankunftszeit am Gate sind die Gegenmittel. Noch schwieriger ist die Situation bei ankommenden Passagieren, die nach der Gepäckausgabe den Flughafen möglichst schnell verlassen wollen. Befragungen zu Ankunftsprozessen brauchen deshalb besonders kurze Fragebögen und einen Befragungsort, an dem Passagiere ohnehin kurz innehalten, etwa an Ausgängen zu Taxi, Bahn oder Parkhaus.

Drittens ist die Ablehnungsbereitschaft ungleich verteilt. Geschäftsreisende und Passagiere mit knappem Zeitpuffer lehnen häufiger ab als Urlaubsreisende mit langem Aufenthalt. Die konsequente Einhaltung der Zählregel, eine kurze und klare Ansprache und die Dokumentation von Ablehnungen nach Passagiergruppe machen diese Verzerrung sichtbar und begrenzbar.

6.2 Repräsentativität und seltene Situationen

Repräsentativität ist keine Eigenschaft einer Stichprobe an sich, sondern immer eine Eigenschaft in Bezug auf eine Frage. Eine standardisierte Erhebung kann die Passagierstruktur eines Jahres hervorragend abbilden und trotzdem für die Frage, wie Passagiere einen Gate-Wechsel erleben, nahezu wertlos sein. An einem durchschnittlichen Tag ist nur ein kleiner Teil der Passagiere davon betroffen, sodass selbst eine große Stichprobe nur wenige Fälle enthält.

Für seltene Situationen gilt deshalb: Sie brauchen eine gezielte Stichprobe, keine größere. Statt die Gesamtstichprobe zu vergrößern, wird die Situation selbst zum Auswahlkriterium, etwa indem Interviewer bei einem Gate-Wechsel kurzfristig die betroffenen Passagiere befragen. Eine ergänzende Filterfrage in der standardisierten Erhebung kann parallel den Anteil der Betroffenen über das Jahr erfassen und so die gezielte Studie einordnen.

Ein verwandtes Problem sind Durchschnittswerte über Situationen, die sich stark unterscheiden. Die Zufriedenheit mit der Sicherheitskontrolle über den gesamten Tag gemittelt kann unauffällig sein, obwohl sie in den Spitzenstunden am Morgen deutlich einbricht. Weil in diesen Stunden besonders viele Passagiere die Kontrolle passieren, betrifft das Problem mehr Menschen, als der Tagesdurchschnitt vermuten lässt.

USE CASE Der Tagesdurchschnitt verdeckt die Morgenspitze

Ausgangslage: Die Zufriedenheit mit der Sicherheitskontrolle lag an einem Flughafen im Jahresdurchschnitt im guten Bereich. Gleichzeitig häuften sich Beschwerden über lange Wartezeiten am frühen Morgen.

Methodisches Vorgehen: Die Befragung wurde für die Zeitfenster der ersten Abflugwelle gezielt aufgestockt, sodass diese Stunden getrennt ausgewertet werden konnten. Die Antworten wurden mit den gemessenen Durchlaufzeiten je Viertelstunde verknüpft.

Ergebnis: In der Morgenspitze lag die Zufriedenheit deutlich unter dem Tagesdurchschnitt, und die Wahrnehmungsdifferenz war hier am größten. Passagiere dieser Zeitfenster hatten zugleich den knappsten Zeitpuffer bis zum Abflug.

Konsequenz: Der Flughafen passte die Schichtplanung so an, dass zu Beginn der ersten Abflugwelle alle Kontrollspuren geöffnet sind, und empfahl Passagieren früher Abflüge über App und Buchungsbestätigung einen größeren Zeitpuffer. Die getrennte Auswertung der Spitzenstunden wurde fester Bestandteil der Berichterstattung.

6.3 Internationale Passagiere

An Flughäfen mit internationalem Verkehr ist ein relevanter Teil der Passagiere nicht in deutscher Sprache erreichbar. Wie gut eine Befragung damit umgeht, entscheidet unmittelbar über ihre Aussagekraft. Wer nur auf Deutsch und Englisch befragt, erreicht bestimmte Passagiergruppen systematisch schlechter, und die Stichprobe verschiebt sich, ohne dass dies in den Ergebnissen sichtbar wird.

Sprachversionen und Übersetzung.
Der Fragebogen sollte in allen Sprachen vorliegen, die einen relevanten Anteil am Passagieraufkommen haben. Eine wörtliche Übersetzung reicht dabei nicht aus. Bewährt hat sich ein mehrstufiges Verfahren aus Übersetzung, Prüfung durch eine zweite muttersprachliche Person und Pretest mit Passagieren der jeweiligen Sprachgruppe. Besondere Aufmerksamkeit brauchen Skalenbeschriftungen und flughafenspezifische Begriffe, die in anderen Sprachen unterschiedlich verstanden werden können.

Sprachkompetenz im Feld.

Welche Sprachen gebraucht werden, lässt sich weitgehend planen. Der Flugplan zeigt, an welchem Gate zu welcher Zeit mit welchen Sprachgruppen zu rechnen ist, und die Einsatzplanung kann Interviewer mit passenden Sprachkenntnissen diesen Zeitfenstern zuordnen. Im Mixed-Mode ist der CASI-Teil im Vorteil, weil er in vielen Sprachversionen parallel angeboten werden kann.

Ansprache.
Die Felderfahrung zeigt, dass Passagiere unterschiedlich auf eine Befragung reagieren. Manche erwarten eine kurze, sachliche Erklärung, andere wollen zunächst wissen, wer die Studie durchführt und was mit ihren Angaben geschieht. Manche lehnen direkt ab, andere zeigen Zurückhaltung indirekt, etwa durch ausweichende Antworten oder eine höfliche Zustimmung ohne echte Bereitschaft. Diese Unterschiede hängen mit Sprache und kulturellem Hintergrund zusammen, ebenso aber mit Reiseanlass, Reiseerfahrung und der konkreten Situation. Pauschale Zuschreibungen nach Herkunft helfen deshalb wenig. Entscheidend ist, dass Interviewer geschult sind, die Signale der einzelnen Person zu lesen, Tempo und Direktheit der Ansprache anzupassen und indirekte Ablehnung zu erkennen. Ob das gelingt, wird gemessen und nicht angenommen. Werden Ablehnungen und Abbrüche je Sprachgruppe dokumentiert, zeigt sich früh, wo der Anteil erfolgreich geführter Interviews, die sogenannte Ausschöpfung, abfällt.

Vergleichbarkeit von Bewertungen.

Auch wenn alle Gruppen gut erreicht werden, bleibt eine methodische Herausforderung, denn Menschen nutzen Bewertungsskalen unterschiedlich. Die Forschung zu Antwortstilen zeigt, dass Befragte je nach Hintergrund eher zu den Extremen oder eher zur Mitte einer Skala neigen oder Aussagen eher zustimmen (Baumgartner und Steenkamp 2001). Ein niedrigerer Zufriedenheitswert einer Passagiergruppe ist deshalb nicht automatisch ein Hinweis auf schlechtere Erfahrungen. Belastbarer als der Vergleich absoluter Werte zwischen Gruppen ist der Vergleich innerhalb einer Gruppe, etwa zwischen Touchpoints oder über Wellen, sowie die Frage nach konkreten Erlebnissen statt allgemeiner Urteile.

7. Fazit: Von der Befragung zur Entscheidung

Fluggastbefragungen sind mehr als ein Stimmungsbarometer. Richtig aufgesetzt, schließen sie die Lücke zwischen dem, was am Flughafen gemessen wird, und dem, was Passagiere erleben. Diese Lücke ist keine Randgröße. Sie entscheidet darüber, ob Investitionen in Prozesse, Flächen und Leitsysteme bei den Passagieren ankommen, ob Non-Aviation-Angebote ihr Potenzial ausschöpfen und ob ein Flughafen im Wettbewerb um Passagiere und Airlines überzeugende Argumente hat.

Die Beispiele in diesem Whitepaper zeigen ein wiederkehrendes Muster. Der Erkenntniswert einer Befragung entsteht selten aus einem einzelnen Durchschnittswert. Er entsteht aus der richtigen Stichprobe, aus der Segmentierung nach den relevanten Gruppen, aus der Verknüpfung mit den operativen Kennzahlen und aus Analysen, die zeigen, welche Faktoren tatsächlich zählen. Ob die Morgenspitze getrennt betrachtet wird, ob Nichtnutzer befragt werden oder ob ein Moduswechsel sauber begleitet wird, entscheidet oft mehr über die Qualität einer Entscheidung als die Zahl der Interviews.

Was eine erfolgreiche Fluggastbefragung ausmacht

1. Die richtige Frage und der passende Studientyp.
Wer die Struktur seiner Passagiere kennen und sich mit anderen Flughäfen vergleichen will, braucht eine standardisierte Erhebung. Wer einen konkreten Prozess oder ein Angebot verbessern will, braucht eine standortspezifische Studie. Beides lässt sich verbinden, aber nicht verwechseln.

2. Eine Stichprobe, die zur Frage passt.

Entscheidend ist nicht die Gesamtzahl der Interviews, sondern die Zahl in der Gruppe, über die eine Aussage getroffen werden soll, und die konsequente Auswahl nach festen Regeln.

3. Feldkompetenz am Flughafen.

Sicherheitsauflagen, Zeitdruck, Gate-Wechsel und internationale Passagiere verlangen erfahrene Teams, die den Flughafenbetrieb kennen und die Qualität der Daten auch unter wechselnden Bedingungen sichern.

4. Interpretation im Zusammenspiel mit den operativen Kennzahlen.
Befragungsdaten entfalten ihren vollen Wert erst in der Verbindung mit dem, was der Flughafen ohnehin misst. Die eigentliche Leistung liegt in der Übersetzung beider Perspektiven in konkrete Handlungsempfehlungen.

Quellen

ACI World, Airports Council International (o. J.): Airport Service Quality (ASQ). Online verfügbar unter https://aci.aero/programs-and-services/asq/ (abgerufen am 01.10.2026).

Baumgartner, H. und Steenkamp, J.-B. E. M. (2001): Response Styles in Marketing Research: A Cross-National Investigation. Journal of Marketing Research, 38(2), S. 143-156.

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J. und Messer, B. L. (2009): Response Rate and Measurement Differences in Mixed-Mode Surveys Using Mail, Telephone, Interactive Voice Response (IVR) and the Internet. Social Science Research, 38(1), S. 1-18.

Flughafenverband ADV (2025): Umfassende repräsentative Fluggastbefragung: Neue Trends im Luftverkehr zeichnen sich ab. Pressemitteilung vom 13.05.2025. Online verfügbar unter https://www.adv.aero/umfassende-repraesentative-fluggastbefragung-neue-trends-im-luftverkehr-zeichnen-sich-ab/ (abgerufen am 01.10.2026).

Flughafenverband ADV (o. J.): Airport Travel Survey. Repräsentative Fluggastbefragung der ADV. Online verfügbar unter https://www.adv.aero/fachbereiche/markt-und-verkehr/airport-travel-survey-repraesentative-fluggastbefragung-der-adv/ (abgerufen am 01.10.2026).

Johnson, J. W. (2000): A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression. Multivariate Behavioral Research, 35(1), S. 1-19.

Kish, L. (1965): Survey Sampling. New York: Wiley.

Maister, D. H. (1985): The Psychology of Waiting Lines. In: Czepiel, J. A., Solomon, M. R. und Surprenant, C. F. (Hrsg.): The Service Encounter: Managing Employee/Customer Interaction in Service Businesses. Lexington, MA: D. C. Heath, Lexington Books, S. 113-123.

Reichheld, F. F. (2003): The One Number You Need to Grow. Harvard Business Review, 81(12), S. 46-54.

Stand: Oktober 2026
Autorinnen: Kordula Later, Sandra Baethge

IWD market research GmbH
Hasselbachplatz 3
39104 Magdeburg
Deutschland

www.iwd-marketresearch.de