

Datenqualität und Information im Verkehrswesen

Max Herry

(Dr. Max HERRY, Verkehrsplanungsbüro, 1040 Wien, Argentinierstr. 21, Tel.: +43-1-504 12 58, Fax: +43-1-504 35 36, e-mail: herry@magnet.at)

GLIEDERUNG:

1. Einleitung
2. Grundlagen
 - 2.1 Information
 - 2.1.1 Was ist Information?
 - 2.1.2 Information und Kommunikation
 - 2.1.3 Daten und Information
 - 2.2 Datenqualität
 - 2.2.1 Was ist Datenqualität?
 - 2.2.2 Messen (von Daten)
3. Grundlegende Zusammenhänge
 - 3.1 Datengewinnung und Information
 - 3.1.1 Empirische Daten
 - 3.1.2 Analytische Daten
 - 3.2 Information und Datengenauigkeit
 - 3.2.1 Datengenauigkeit von empirischen Daten
 - 3.2.2 Datengenauigkeit von analytischen Daten
 - 3.2.3 Wie genau ist „genau genug“?
 - 3.3 Verarbeitung von unvollständigen oder fehlenden Informationen
 - 3.3.1 Arbeiten mit unvollständigen Informationen
 - 3.3.2 Non-Reported-Topics-Analyse
 - 3.3.3 Arbeiten mit fehlenden Informationen
 - 3.3.4 Non-Response-Analyse
 - 3.4 Daten(-Interpretation) als Information
 - 3.4.1 Informationsgehalt - Entropie
 - 3.4.2 Perzeption und Wirklichkeit
 - 3.4.3 Datenmengen und Information
 - 3.4.4 Wie lügt man mit Statistik
 - 3.5 Informationen über Daten - Datenzugang
 - 3.5.1 Informationslust - „Wissen ist Macht!“
 - 3.5.2 Informationsfrust - „Was ich nicht weiß, macht mich nicht heiß!“
 - 3.5.3 Informationsschutz - Datenschutz
4. Weitere Zusammenhänge zwischen Datenqualität und Information
 - 4.1 Daten und Informationsgewinn
 - 4.1.1 Erkenntnis-(Modelle) und Information
 - 4.1.2 Verkehrsmodelle und Information
 - 4.1.3 Komplexität von Modellen
 - 4.2 Informationsrückkopplung
 - 4.2.1 Kybernetischer Regelkreis
 - 4.2.2 Modelle mit Informationsrückkopplung
 - 4.3 Information und Entscheidungsverhalten
 - 4.3.1 Entscheidungsverhalten
 - 4.3.2 Information und Kooperation
 - 4.4 Information und Bestimmtheit
 - 4.4.1 Simulationen
 - 4.4.2 Information und Chaos
 - 4.5 Formen von Informationen - Datenformen
 - 4.5.1 Individual-Information und Kollektiv-Information
 - 4.5.2 Aggregation und Disaggregation von Daten
5. Schlußfolgerungen
6. Literatur

1. EINLEITUNG

Datenqualität und Information ist ein unerschöpfliches Thema, wenn man es allgemein und umfassend darlegen möchte. In diesem Manuskript werden nur einige wesentliche Punkte zum Vortrag angerissen und in einer Kurzfassung ausgewiesen. Das Thema ist vor allem deshalb wichtig, weil

- einerseits, die Information zum wesentlichen Bestandteil unseres heutigen Lebens geworden ist (Informationsgesellschaft) und auch eine der Grundlagen für die „Bestimmung“ und Gestaltung **unserer** Zukunft sein wird,
- andererseits, aber gerade dazu - und das vor allem in der Verkehrsplanung - nicht nur Daten an sich, sondern Daten von hoher Qualität gebraucht werden.

Die Darstellung konzentriert sich auf die Belange der Verkehrsforschung und -planung und beinhaltet nicht nur Statements, sondern vor allem auch Anregungen zur Verbesserung der Datenqualitäten und Informationen.

2. GRUNDLAGEN

Die Grundlagen zum gegenständlichen Vortragsthema bilden die Begriffe „Information“ und „Datenqualität“.



2.1. Information

Zum ersten Bereich bringen wir, was Information eigentlich ist sowie den engen Zusammenhang zwischen Information und Kommunikation.

2.1.1. Was ist Information?

Es ist gar nicht mal so einfach, bei den Informations-Experten - das sind nicht nur die Informatiker - oder in der entsprechenden Literatur klare Informationen zum Begriff „Information“ zu erhalten. Zum großen Teil drücken sich die Fachleute um diesen Bereich und gehen gleich zur Informationstheorie bzw. zur Kommunikationswissenschaft über.

Das ist oft so mit grundlegenden Begriffen: So wird in der Verkehrsforschung oder -planung der fundamentale Begriff des Verkehrs entweder überhaupt nicht oder nur sehr „mystisch“ erklärt.

Nach SHANNON¹ ist die Information I als eine Funktion von zwei Variablen definiert:

$$I(A;B) \stackrel{\text{def}}{=} \log_b \frac{P(A|B)}{P(A)}$$

Es wird vorausgesetzt, daß $P(A) \neq 0$ und $P(B) \neq 0$ gilt, wobei $P(A)$ die (Eintritts)-Wahrscheinlichkeit des Ereignisses A ist.

Auch JOHANNESSON weist darauf hin², daß die Information eine **Funktion zweier Variablen** ist. Fragt man einen Informationstheoretiker, wieviel Information eine Nachricht enthält, so wird er fragen: „Worüber?“. FORNEY meint jedoch, daß dieser Punkt in fast allen populärwissenschaftlichen Darstellungen der Informationstheorie fehlt, die normalerweise nur die einvariablige Funktion der Entropie (siehe Kapitel 3.4.1) diskutieren. Alle wichtigen Resultate in der Informationstheorie seien mit dem Informationsbegriff verbunden und nicht mit dem Entropiebegriff!³

Informationen können nun im Verkehrsbereich

- **mengenmäßiger** Natur sein, wie zum Beispiel
 - Verkehrsbelastungen oder -leistungen
 - oder Straßenmerkmale,
- **zeitlicher** Natur sein, wie zum Beispiel
 - Reisezeiten
 - oder Wegedauer,
- **mengen- und zeitmäßiger** Natur sein, wie zum Beispiel
 - Ganglinien (Tages-, Wochen- oder Jahresganglinien)
 - oder Dauerlinien⁴
- **räumlicher** Natur sein, wie zum Beispiel
 - Verkehrsnetze,
 - Verkehrsströme
 - oder Verkehrsspannungen

sein, die ihrerseits in vielfältiger Weise weiter differenziert werden können,

- zum Beispiel nach Wegezwecken, Gütergruppen, Verkehrsarten, Fahrzeugarten, Entfernungsbereichen, Zeitbereichen usw.

Sie können aber auch

- Einstellungen,
- Meinungen und
- Akzeptanzen

der VerkehrsteilnehmerInnen und auch Nicht-VerkehrsteilnehmerInnen beinhalten.

In der Abbildung 1 sind zum Beispiel Aspekte und Indikatoren (Informationen) zur Mobilität dargestellt.

Eine Information kann auch einen Strukturzusammenhang wiedergeben, wie in der Abbildung 2 gezeigt wird.

Wichtig ist nun,

- nicht nur zu wissen, auf **welchen** Informationen die davon erzeugten Daten aufbauen, sondern auch - und vor allem -
- welche dieser Informationen **unvollständig** sind,
- und welche Informationen dazu **überhaupt nicht** zur Verfügung stehen.

¹ SHANNON C.E.: A Mathematical Theory of Communication. Bell Sys Tech J. 27:379-423 (Part I), 1948

² JOHANNESSON, R.: Informationstheorie. Verlag Addison-Wesley, Lund 1988

³ FORNEY, G. D., Jr.: Information Theory, course notes for EE-376. Stanford University, 1972

⁴ HERRY M.: Methode und mathematisches Grundmodell zu Berechnungen im Mengengerüst von Entscheidungshilfen. Diss. An der TU Wien, Wien 1982

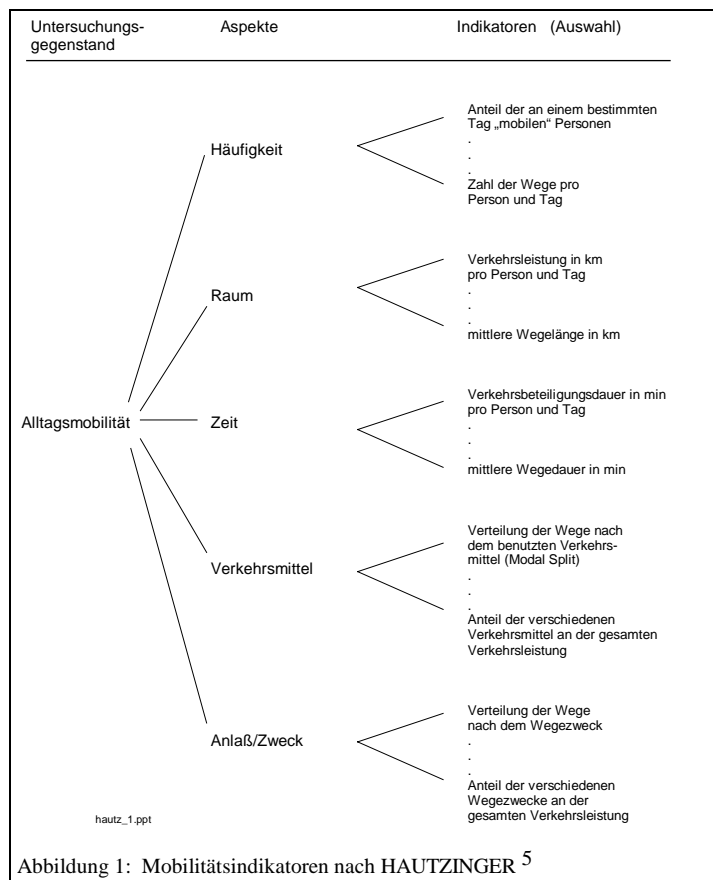


Abbildung 1: Mobilitätsindikatoren nach HAUTZINGER⁵

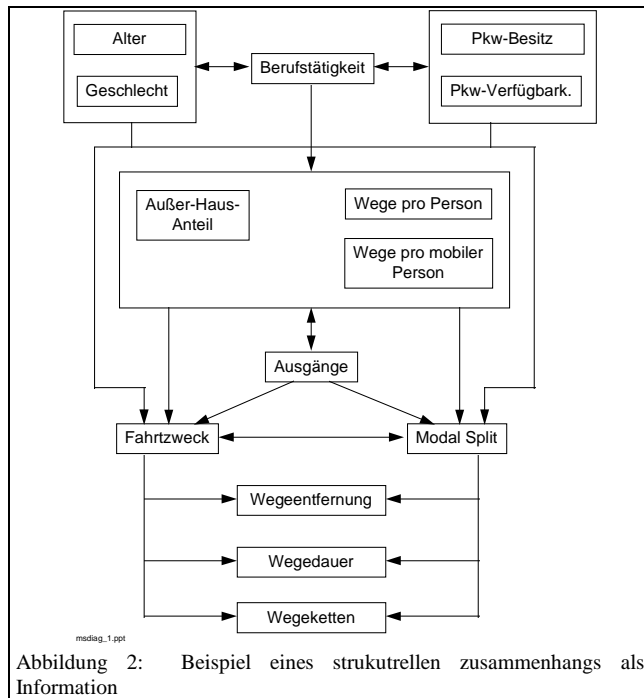
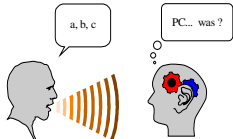


Abbildung 2: Beispiel eines strukturellen Zusammenhangs als Information

Der zweite Bereich bedingt eine Non-Reported-Analyse⁶, der dritte eine Non-Response-Analyse⁷, Verfahren, die leider allzu oft bei der Datengewinnung außer acht gelassen oder zumindestens unterschätzt werden (siehe dazu Kapitel 3.3.2 und 3.3.4).



2.1.2. Information und Kommunikation

Eng verbunden mit dem Begriff der Information ist die Kommunikation. Der Begriff der Kommunikation, aus dem lat. übersetzt, steht für Mitteilung, Verbindung oder auch Verkehr. So gesehen bedeutet Kommunikation weit mehr als umgangssprachlich angenommen. Nicht umsonst unterscheidet

man Kommunikation von Konversation:

- Oft spricht man davon, daß Leute miteinander reden, ohne sich dabei etwas zu sagen - sie kommunizieren nicht,
- oder umgekehrt, daß sich zwei verstehen, ohne ein Wort gewechselt zu haben - nonverbale Kommunikation.

Unter **Kommunikation** soll folglich

- jede Art der annähernden Interaktion zwischen zwei oder mehreren am Kommunikationsprozeß teilnehmenden Elementen verstanden werden.
- Die Interaktion ist jene bewußte oder unbewußte Handlung, die Informationen an ein anderes Element weitergibt und umgekehrt von anderen Elementen Informationen aufnimmt.
- Ein Kommunikationselement kann jedes Objekt, das die Fähigkeit des Sendens, Empfangens und der Informationsverarbeitung und -bearbeitung besitzt, sein⁸.

Das Forschungsgebiet der Informationstheorie, die Grundlage der gesamten Telekommunikation, wurde 1948 von Claude SHANNON mit seinem aufsehenerregenden Aufsatz: „A Mathematical Theory of Communication“ begründet. Als der Aufsatz ein Jahr später in Buchform erschien, wurde der Titel bemerkenswerterweise in „The Mathematical Theory of Communication“ geändert.

Von großen wissenschaftlichen Entdeckungen wird im nachhinein oft gesagt, daß die Zeit für diesen Durchbruch reif war. Nicht so bei der Informationstheorie: Shannons Ergebnisse waren so revolutionär, daß viele der damaligen Kommunikationsspezialisten deren Bedeutung nicht verstanden: Erst nach und nach wurde deutlich, daß Shannon eine ganz neue Wissenschaft begründet hatte⁹.

⁵ HAUTZINGER H. et al.: Mobilität - Ursachen, Meinungen, Gestaltbarkeit. Studie im Auftrag des VDA, der BAG und des ADAC, Heilbronn 1994

⁶ BRÖG W., ERL E., MEYBURG A.H.: Problems of Non-reported Trips in Surveys of Non-home Activity Patterns. 61st Annual Meeting of the Transportation Research Board, Washington, D.C., January, 18-22, 1982

⁷ HERRY M., SAMMER G.: Gewichtung und Hochrechnung der Österreichischen Verkehrsverhaltenshebung. Im Auftrag des Bundesministeriums für Wissenschaft, Verkehr und Kunst, Wien 1997

⁸ MIZEROVSKY H.: Kommunikation im Ernstfall. Wien - München - Zürich

⁹ MASSEY J.L.: Information Theory: the Copernican System of Communications. IEEE Commun Mag. 22:26-28, 1984

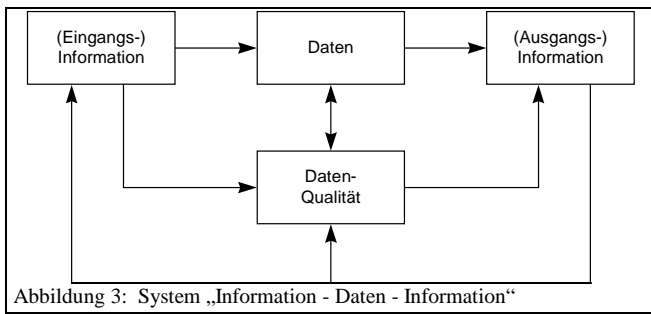


Abbildung 3: System „Information - Daten - Information“

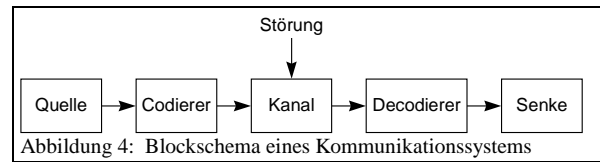


Abbildung 4: Blockschema eines Kommunikationssystems

2.1.3. Daten und Information

Um Daten zu gewinnen (erstellen), braucht man einerseits bestimmte Informationen (Eingangsinformationen). Die (erstellten) Daten liefern ihrerseits wiederum Informationen (Ausgangsinformationen) durch Dateninterpretation und -weiterverarbeitung. Die Eingangsinformationen bestimmen aber nicht nur die Daten selbst, sondern mit ihnen auch ihre (Daten-)Qualität. Je höher nun diese Qualität ist, um so größer ist der Informationsgewinn aus diesen Daten.

Bei unzureichender Qualität der Daten bzw. der Ausgangsinformationen sollten entsprechende Rückkopplungen in diesem System (siehe Abbildung 3) stattfinden.

Der Prozeß der Rückkopplung spielt im Bereich der Informationstheorie und -praxis eine enorm wichtige Rolle. Durch sie konnten erst solche bedeutende Bereiche, wie

- (kybernetische) Regelkreise (siehe Kapitel 0) oder
- auch die Chaos-Theorie (siehe Kapitel 4.4.2)

entwickelt werden.

Die Beziehung „Daten - Information“ erzeugt also vier wesentliche Zusammenhangsbereiche:

- Information --> Daten (*Eingangsinformationen* für Datenerstellung)
- Daten als Information (*Dateninhalt*)
- Daten --> Information (*Informationsgewinnung* aus Daten(-Weiterverarbeitung) - *Ausgangsinformation*)
- Information über die Daten (*Informationszugang*)

Siehe dazu auch die Abbildung 5.

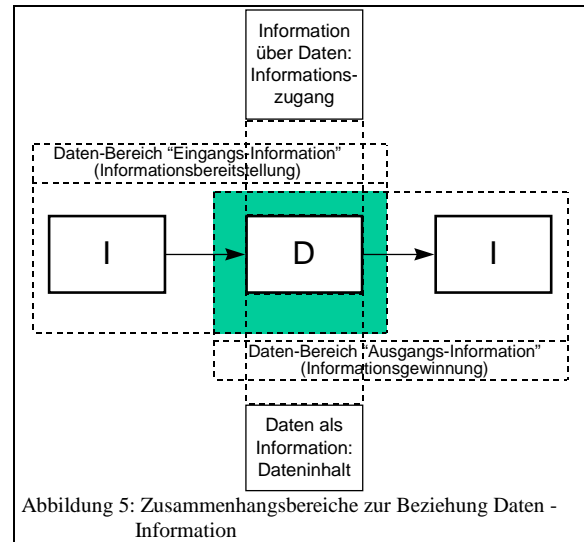


Abbildung 5: Zusammenhangsbereiche zur Beziehung Daten - Information

Datenqualität



Der zweite wichtige Bereich zum Vortragsthema ist die Datenqualität. Wir klären zunächst, was wir darunter verstehen und gehen anschließend in diesem Kapitel noch auf den Aspekt des Messens (von Daten) ein.

2.2.1. Was ist Datenqualität?

Datenqualität ist eher ein umgangssprachlicher Begriff. In der Verkehrsforschung und -planung wird darunter vor allem

- die **Datengenauigkeit** (*Abbildungsschärfe*) bzw.
- die Möglichkeit, die Daten mit einem hohen **Informationsgewinn** weiterverarbeiten (oder wenigstens interpretieren) zu können,

verstanden.

Wir gehen auf diese Bereiche in den Kapiteln 3.1 und 3.2 weiter ein.

Datenqualität hat damit etwas mit:

- Stichprobenfehlern,
- Zufallsfehlern,
- inhaltlichen Fehlern und
- Erfassungsfehlern

zu tun.

Im allgemeinen - in der Diskussion um diesen fundamentalen Teil bei der Datenbehandlung - wie im besonderen - bei einer konkreten Datenerstellung - wird jedoch dieser Aspekt meist nur mit „vollem Mund“, aber dafür mit „leerem Methodensack“ begleitet.

2.2.2. Messen (von Daten)

Datenqualität hat offensichtlich etwas mit Messen zu tun: Ein „gutes“ Messen erzeugt eine hohe Datenqualität, ein „schlechtes“ hingegen vermindert diesen Anspruch. Oft wird - insbesondere in der klassischen Testtheorie (KTT) - die Frage erhoben, ob man für die mit einem bestimmten Verfahren erhobenen Daten überhaupt den Anspruch erheben darf, etwas zu *messen*¹⁰. Die

¹⁰ STEYER R., EID M.: Messen und Testen. Springer-Verlag Berlin Heidelberg New York, 1993

Beantwortung dieser Frage bedingt die Präzisierung, was man unter „Messen“ verstehen will. Frei nach STEYR / EID ¹¹ bedeutet dies die Postulierung bestimmter (empirischer) Gesetze mit einem Meßmodell (siehe Kapitel 3.2.1.1), aus deren Gültigkeit die Existenz einer theoretischen Größe (mit bestimmten Eigenschaften) logisch abgeleitet werden kann. Dabei entsteht das **Meßfehler-Problem**, das folgendermaßen dargestellt werden kann:

$$Y_i(x) = \tau(x) + \epsilon_i(x) \tag{F1}$$

wobei:

Y der (zum i-ten Male) gemessene Wert einer Größe x,
 τ(x) der „wahre“ Wert von x und
 ε_i(x) der Meßfehler der (i-ten) Messung von x
 ist, und der Erwartungswert von Y gleich τ ist.

Dieser Sachverhalt stellt für unsere weiteren Betrachtungen einen wesentlichen Ausgangspunkt dar.

3. GRUNDLEGENDE ZUSAMMENHÄNGE

Bei den grundlegenden Zusammenhängen geht es vor allem um die Beziehungen von Information und:

- Datengewinnung,
- Datengenauigkeit,
- Verarbeitung von unvollständigen oder fehlenden Informationen,
- Dateninterpretation und
- Informationen über Daten.

3.1. Datengewinnung und Information

Verkehrsdaten können

- **erhoben** - **empirische** Daten -
 - und/oder **berechnet** - **analytische** Daten -
- werden.

3.1.1. Empirische Daten

Empirische Daten werden **erhoben**.

Die grundsätzlichen Zusammenhänge bei der Erhebung von Verkehrsdaten sind in der Abbildung 6 ersichtlich.

Die Erhebung kann in Form

- einer **Zählung** (passive Erhebungsmethode) oder
- einer **Befragung** (aktive Erhebungsmethode)

durchgeführt werden.

Bei den Befragungen kommen vor allem folgende Ansätze zur Anwendung ¹¹:

- Revealed-Preferences-Verfahren (RP-Verfahren),
- interaktive Meßverfahren und
- Stated-Preferences-Verfahren (SP-Verfahren).

RP- oder Revealed-Preferences-Verfahren erheben **realisiertes** Verkehrsverhalten (aus der Vergangenheit):

<ul style="list-style-type: none"> • Vorteile: <ul style="list-style-type: none"> ◦ kein antizipiertes Verhalten, ◦ weites Spektrum von individuellen Randbedingungen, ◦ genaue Beschreibung von Marktsegmenten möglich, ◦ ausgearbeitete Korrektur- und Gewichtungungsverfahren, ◦ standardisierte Fragebogen vorhanden, ◦ ausgearbeitete Analyseverfahren und ◦ gute Abschätzung der systematischen und zufälligen Fehler, 	<ul style="list-style-type: none"> • Nachteile: <ul style="list-style-type: none"> ◦ nur vorhandene Maßnahmen und Situationen erhebbar/befragbar, ◦ subjektive Einflußfaktoren nur gering, ◦ Einflußgrößen müssen bekannt sein, ◦ qualitative Einflußgrößen nur gering, ◦ großer Stichprobenumfang erforderlich, ◦ Erhebungsdauer, ◦ Stichprobenziehung und Hochrechnung erfordert Erfahrung, ◦ Non-Response-Analyse erforderlich, ◦ Non-Reported-Trips-Analyse erforderlich und ◦ Korrelationen zwischen Einflußgrößen müssen beachtet werden.
--	--

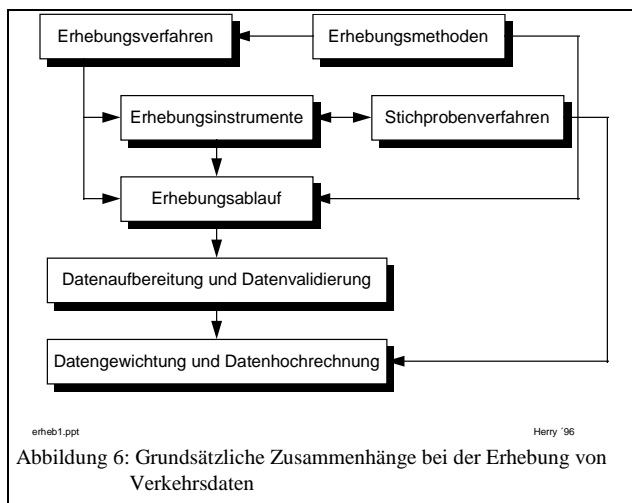


Abbildung 6: Grundsätzliche Zusammenhänge bei der Erhebung von Verkehrsdaten

¹¹ HERRY M.: Mobilität von Personen und Gütern. Vorlesung an der TU Wien, Wien 1996

Interaktive Verfahren werden durch folgende Bereiche gekennzeichnet:

- interaktives Gespräch,
- Durchführung von Hinterfragungen und
- Reorganisation des getätigten Verkehrsverhaltens.

Zu den Vor- und Nachteilen von interaktiven Verfahren ist folgendes anzuführen:

<ul style="list-style-type: none"> • Vorteile: <ul style="list-style-type: none"> ◦ neue Situationen abfragbar, ◦ Berücksichtigung "schwacher" qualitativer Einflußgrößen, ◦ Situationsanalyse, ◦ Hinterfragung möglich, ◦ Reorganisation der untersuchungsrelevanten Aktivitäten, ◦ hohes Maß an Zuverlässigkeit der Antworten, ◦ Freiheit der zu testenden Maßnahmen, ◦ Lernfähigkeit des Systems, ◦ Einsatz von Telefon und Computer möglich und ◦ Einbeziehung von Familie bzw. eines größeren Befragungsumfeldes möglich. 	<ul style="list-style-type: none"> • Nachteile: <ul style="list-style-type: none"> ◦ Analyse mehrerer Marktsegmente nur schwer möglich, ◦ Antwort vs. Verhalten, ◦ Erfahrung bei Gestaltung erforderlich, ◦ keine Standardisierung der Interaktionen, ◦ gute Interviewer erforderlich, ◦ Auswertungsverfahren unbefriedigend, ◦ Stichprobenwahl vs. Anteil der "Reagierer", ◦ nur wenige Maßnahmen möglich und ◦ relative Bedeutung einzelner Einflußgrößen schwer erkennbar.
---	---

Die Methoden der SP- oder **Stated-Preferences-Verfahren**¹², die auch **Conjoint Analysis**, **Contingent Valuation** oder **Direct Utility Assessment** genannt werden, ermitteln mögliche Verhaltensreaktionen der Befragten durch Vorgabe einer Reihe von Entscheidungssituationen, die verschiedene Alternativen enthalten. Diese Situationen können sowohl real als auch hypothetisch sein. Abbildung 7 zeigt dazu ein Beispiel.

					Wahl
Pkw	20 Min. Fahrzeit	7 Min Suchzeit für den Parkplatz im Parkhaus	4 Min. Fußweg zum Ziel	Parkgebühr 30 öS	
ÖV	25 Min. Fahrzeit ohne Umsteigen	10 Min. Wartezeit	8 Min . Fußweg zum Ziel	Fahrpreis 20 ÖS	
Pkw	10 Min. Fahrzeit	2 Min Suchzeit für den Parkplatz im Parkhaus	1 Min. Fußweg zum Ziel	Parkgebühr 10 öS	
ÖV	14 Min. Fahrzeit ohne Umsteigen	3 Min. Wartezeit	8 Min . Fußweg zum Ziel	Fahrpreis 15 ÖS	
Pkw	20 Min. Fahrzeit	2 Min Suchzeit für den Parkplatz in der Kurzparkzone	2 Min. Fußweg zum Ziel	Parkgebühr 18 öS	
ÖV	14 Min. Fahrzeit mit einmal Umsteigen	10 Min. Wartezeit	5 Min . Fußweg zum Ziel	Fahrpreis 10 ÖS	
Pkw	13 Min. Fahrzeit	7 Min Suchzeit für den Parkplatz in der Kurzparkzone	1 Min. Fußweg zum Ziel	Parkgebühr 10 öS	
ÖV	18 Min. Fahrzeit mit einmal Umsteigen	10 Min. Wartezeit	3 Min . Fußweg zum Ziel	Fahrpreis 10 ÖS	
Pkw	20 Min. Fahrzeit	7 Min Suchzeit für den Parkplatz im Parkhaus	4 Min. Fußweg zum Ziel	Parkgebühr 30 öS	
ÖV	14 Min. Fahrzeit mit einmal Umsteigen	6 Min. Wartezeit	3 Min . Fußweg zum Ziel	Fahrpreis 10 ÖS	
Pkw	13 Min. Fahrzeit	2 Min Suchzeit für den Parkplatz am Straßenrand	4 Min. Fußweg zum Ziel	Parkgebühr keine	
ÖV	18 Min. Fahrzeit mit einmal Umsteigen	3 Min. Wartezeit	5 Min . Fußweg zum Ziel	Fahrpreis 10 ÖS	
Pkw	20 Min. Fahrzeit	2 Min Suchzeit für den Parkplatz in der Kurzparkzone	2 Min. Fußweg zum Ziel	Parkgebühr 18 öS	
ÖV	18 Min. Fahrzeit ohne Umsteigen	6 Min. Wartezeit	8 Min . Fußweg zum Ziel	Fahrpreis 15 ÖS	
Pkw	10 Min. Fahrzeit	7 Min Suchzeit für den Parkplatz im Parkhaus	1 Min. Fußweg zum Ziel	Parkgebühr 18 öS	
ÖV	25 Min. Fahrzeit ohne Umsteigen	3 Min. Wartezeit	5 Min . Fußweg zum Ziel	Fahrpreis 15 ÖS	

Bei diesem Beispiel ist allerdings nicht zu übersehen, daß die Grenzen einer validen Informationsübertragung vom Befragten erreicht sind.

Abbildung 7: Beispiel einer SP-Befragung¹³

¹² AXHAUSEN K., BOGNER W., HERRY M., VERRON H., VOLKMAR H., WICHMANN W., ZUMKELLER D.: Merkblatt zur Messung von Präferenzstrukturen - Methoden der „Stated Preferences“. 1996

¹³ HERRY M., AXHAUSEN K.: SP-Befragung der Wiener Bevölkerung zur Verkehrsverhaltenshebung Wien 1991. Im Auftrag des Magistrats der Stadt Wien, Wien 1992

Zu den Vor- und Nachteilen von SP-Verfahren ist folgendes anzuführen ¹⁴:

<ul style="list-style-type: none"> ● Vorteile: <ul style="list-style-type: none"> ○ systematische Untersuchung des relevanten Teils des Untersuchungsraumes, ○ "deutlich neue" Maßnahmen einbindbar, ○ Berücksichtigung des Einflusses schwacher Einflußgrößen möglich, ○ Hypothesenformulierung, ○ Anwendung von Versuchsplänen, ○ ausgereifte Analysemethoden zur statistischen Zuverlässigkeit der Ergebnisse, ○ Kombination mit Status-Quo-Befragungen und interaktiven Meßverfahren möglich, ○ Wirkung der Einflußgrößen ist unabhängig voneinander meßbar, ○ spezifisches Untersuchungsdesign möglich, ○ mehrere Beobachtungen pro Probanden, ○ Ergebnisse schnell verfügbar, ○ Verwendung standardisierter Fragebögen, ○ Einsatz von Computern möglich und wünschenswert. 	<ul style="list-style-type: none"> ● Nachteile: <ul style="list-style-type: none"> ○ Bestimmung der relevanten Einflußgrößen notwendig, ○ Annahmen über zu berücksichtigende Wechselwirkungen erforderlich, ○ Antwort <-> Verhalten, ○ Hypothesenformulierung notwendig, ○ Entwurf von Versuchsplänen erforderlich, ○ spezielle methodische Erfahrungen bei der Gestaltung erforderlich, ○ statistische Probleme bei der Auswertung, ○ Schätzung von absoluten Größen problematisch, ○ Stabilität der Ergebnisse mitunter nicht einfach zu erreichen.
--	---

Abgesehen von dieser Differenzierung können auch vielfältige andere Unterscheidungen getroffen werden ¹⁵.

3.1.2. Analytische Daten

Analytische Daten werden (mit Hilfe von Modellen) **berechnet** ¹⁷.

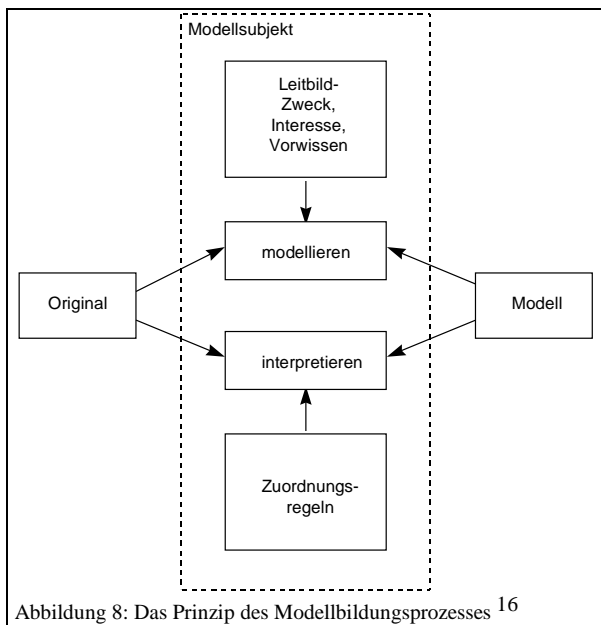


Abbildung 8: Das Prinzip des Modellbildungsprozesses ¹⁶

Abbildung 9 zeigt den Modellbildungsprozeß etwas verfeinert.

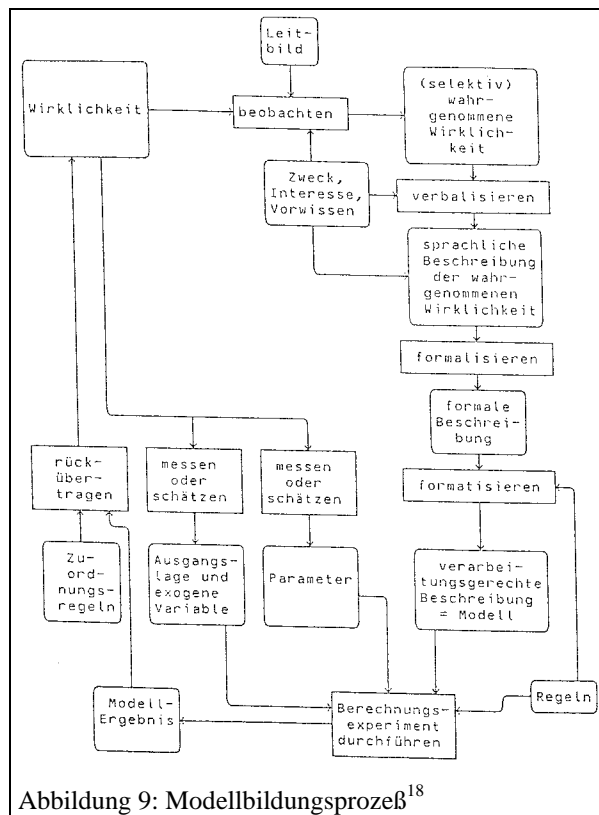


Abbildung 9: Modellbildungsprozeß ¹⁸

¹⁴ HERRY M.: Die methodischen Grenzen der Direkten Nutzenmessung. Kolloquium "Veränderungen im Verkehrsverhalten: Messung und Modellierung mit den Methoden der Direkten Nutzenmessung" am 10. März 1994 in Karlsruhe

¹⁵ AXHAUSEN K.: Erhebungsverfahren im Verkehrswesen. Innsbruck 1996

¹⁶ CRAEMER D.: Mathematisches Modellieren dynamischer Vorgänge. B.G.Teubner, Stuttgart 1985

¹⁷ zum Beispiel: DORFWIRTH J.-R., HERRY M.: Verkehrsmodell Österreich - Methode und mathematisches Grundmodell. Bundesministerium für Bauten und Technik, Straßenforschung, Heft 144, Wien 1980

Die Tabelle 1 gibt einen Überblick über die Arten von solchen Verfahren ¹⁸.

3.2. Information und Datengenauigkeit

Das Thema der Datengenauigkeit und ihr Zusammenhang mit dem Bereich der Information ist eines der **zentralen** Aspekte in der Verkehrsforschung, aber auch Verkehrsplanung. Wir haben dabei wiederum zwischen empirischen und analytischen Daten zu unterscheiden.

3.2.1. Datengenauigkeit von empirischen Daten

Genauigkeit hat etwas mit **Messen** zu tun (vergl. Dazu auch mit Kapitel 2.2.2). Deshalb zunächst einiges zu den Meßmodellen.

3.2.1.1. Messmodelle

In der empirischen Wissenschaft beinhalten Meßmodelle die logische Struktur des theoretischen Begriffs und seine Verknüpfung mit den empirischen Begriffen.

In der folgenden Tabelle sind die wichtigsten Ziele von Meßmodellen aufgelistet ¹⁸.

Die wichtigsten Ziele von Meßmodellen	
<i>Überbrückung von Theorie und Empirie</i>	Wie kann man theoretische Begriffe (z.B. Ängstlichkeit) und empirische Begriffe (Antwortmuster in einem Test) miteinander verbinden?
<i>Explikation der logischen Struktur eines Begriffs</i>	In welchem Kontext ist die Verwendung des Begriffs überhaupt erlaubt? Handelt es sich um einen klassifikatorischen, einen komparatorischen oder einen metrischen Begriff
<i>Einbeziehung des Meßfehlers in das Modell</i>	Wie wirken sich Meßfehler auf die beobachteten Werte aus? Wie stark ist ihr Einfluß, d.h. wie zuverlässig sind die beobachteten Meßwerte?
<i>Einbeziehung situationaler Effekte in das Modell</i>	Wie wirken sich situationale Effekte auf die beobachteten Werte aus? Wie stark ist ihr Einfluß, d.h. wie stabil sind die beobachteten Meßwerte?

Tabelle 2: Ziele von Messmodellen

stochastisch oder deterministisch	unter Benutzung von Zufallsvariablen keine Benutzung von Zufallsvariablen; jeder Faktor ist eindeutig bestimmt, sobald die Faktoren bestimmt sind, mit denen er in Beziehung steht
dynamisch oder statisch	die Zeit tritt explizit als Variable auf die Zeit tritt nicht als Variable auf
mikro- oder makro-	Modell enthält auch Individualdaten, z.B. aus Panels Modell enthält nur aggregierte Daten
linear oder nicht-linear	Änderungen in einer Variablen verursachen nur proportionale Änderungen in anderen Variablen nicht proportionale Änderungen; zur Beschreibung einer nicht-linearen Beziehung werden alle Potenzen der Variablen benutzt außer der Potenz 1, d.h. die Variable selbst
diskret oder kontinuierlich	unter Benutzung von Variablen, die sich schrittweise ändern, z.B. in Schritten von ganzen Zahlen unter Benutzung von Variablen, die sich stetig ändern, meist wird weiter vorausgesetzt, daß der Kurvenverlauf glatt, d.h. "differenzierbar" ist
qualitativ oder quantitativ	Variablen werden mit Nominal- und Ordinalskalen gemessen Variablen werden mit Intervall- und Verhältnisskalen gemessen

Tabelle 1: Klassifikation von Modellen

Man unterscheidet deterministische und stochastische Meßmodelle.

3.2.1.2. Fehler und Information

Im Kapitel 2.2.2 wurde bei der Behandlung des Begriffs „Messen“ auf das Meßfehler-Problem hingewiesen (vergleiche dazu auch mit Gleichung F1).

Wir wollen uns dieser Frage etwas ausführlicher widmen.

In den (Stichproben-)Ergebnissen von empirischen Daten gibt es folgende **Meßfehler**:

<ul style="list-style-type: none"> ● stichprobenbedingte Fehler, resultierend aus <ul style="list-style-type: none"> ○ den Verzerrungen der Stichprobe, bedingt durch <ul style="list-style-type: none"> * Mängel hinsichtlich der Erhebungsgesamtheit, * Verzerrung durch die Auswahlmethode, * Verzerrung durch die Auswahltechnik, * Verzerrung durch das Schätzverfahren ○ Zufallsfehler, 	<ul style="list-style-type: none"> ● nicht-stichprobenbedingte Fehler, resultierend aus <ul style="list-style-type: none"> ○ inhaltlichen Fehlern, bedingt durch <ul style="list-style-type: none"> * die Vorbereitungsphase, * die Erhebungsphase, * die Daten-Aufbereitungsphase, ○ Erfassungsfehlern, bedingt durch <ul style="list-style-type: none"> * Mängel der Erfassungsgrundlage oder * Nichtbeantwortung.
---	--

Die **Zufallsfehler** ¹⁹ lassen sich durch die einschlägigen **Standardverfahren** der Fehlerrechnung ²⁰ abschätzen.

¹⁸ STEYER R., EID M.: Messen und Testen. Springer-Verlag Berlin Heidelberg New York, 1993

So ist zum Beispiel der Standardabweichung des Schätzwertes x des Gesamtdurchschnitts X bei geschichteter Auswahl:

$$s_x = \sqrt{\sum_{h=1}^L \left(\frac{N_h}{N}\right)^2 \cdot \frac{s_h^2}{n_h} \cdot \left(1 - \frac{n_h}{N_h}\right)} \quad (F2)$$

- mit:
- L Anzahl der Schichten
 - N_h Anzahl der Einheiten in der h-ten Schicht in der Grundgesamtheit
 - N Gesamtzahl der Einheiten in der Grundgesamtheit
 - n_h Anzahl der Einheiten in der h-ten Schicht in der Stichprobe
 - n Gesamtzahl der Einheiten in der Stichprobe
 - s_h Standardabweichung der Einzelwerte der h-ten Schicht in der Stichprobe

Bei den **systematischen** Fehlern

- ist diese Abschätzung jedoch um ein Vielfaches schwieriger ²¹,
- da in diesem Falle die Prozedur der Informationsübertragung sehr genau nachvollzogen werden muß, was,
 - erstens, nur zum Teil möglich ist,
 - zweitens, überhaupt nur bei wenigen Erhebungsinstitutionen Standard ist.

Aber selbst den Zufallsfehlern wird in der Regel nicht nachgegangen!

Das ist dann aber wirklich ein „sträflicher Leichtsin“.

Die Reduktion von systematischen Fehlern ist ein sehr umfangreiches und hinreichend kompliziertes Thema:

- KRUG / NOURMEY ²² versuchen dazu hinreichend komplexe Antworten zu geben, was zur Folge haben kann, daß man dabei den „Wald vor Bäumen“ nicht mehr sieht.
- KÖLTRINGER „vereinfacht“ das Fehlerschema und unterscheidet zwischen
 - Stichprobenfehlern und
 - Meßfehlern, wobei zu den Stichprobenfehlern
 - die statistischen Zufallsfehler (sampling error),
 - die Nichterfassungsfehler (noncoverage error) und
 - die Ausfallfehler (nonresponse error) zählen ²².
 Die **Meßfehler** hingegen setzen sich bei ihm aus
 - den Interviewfehlern,
 - den Methodenmeßfehlern (Fehler, die sich auf die Datenerhebungsmethode und alle Aspekte der Fragenformulierung beziehen) und
 - den Befragtenfehlern zusammen.

Dabei wird die Interviewer-Variabilität für jedes Meßinstrument einer Befragung nach folgendem Modell geschätzt:

$$\text{Meßwert} = \text{„wahrer“ Wert} + \text{Interviewereffekt} + \text{Zufallsfehler} \quad (F3)$$

Nicht zu unterschätzen ist dabei - insbesondere bei schlecht eingeschulten und schlecht bezahlten Interviewern - ein eventuelles Täuschungsverhalten der Interviewer (als ein Teil des Interviewereffekts): Die folgende Tabelle 3 zeigt Ergebnisse hierzu.

Items	"schon vorgekommen"
lange Erzählungen auf offene Fragen nicht vollständig notieren oder verkoden	21.2%
bei offenen Fragen den Befragten "bremsen", um nicht zu viel notieren zu müssen	17.1%
mehr Kontaktversuche angeben, als tatsächlich gemacht wurden	14.3%
das Thema eines Fragenblocks kurz anschneiden, um die Meinung des Befragten zu erfahren, und den Block später selbst ausfüllen	9.4%
wenn beim ersten Kontaktversuch niemand angetroffen wird, Verweigerung, Tod oder Verzug kodieren	7.7%
unangenehme (intime, delikate) Fragen selbst ausfüllen	7.7%
den Fragebogen nach einem Abbruch des Befragten zu Hause fertig ausfüllen	5.1%
bei Filterfragen fälschlich so ankreuzen, daß der folgende Fragenblock übersprungen werden kann	3.4%
in einem "Blitzinterview" einige wichtige Fakten in Erfahrung bringen und den Fragebogen zu Hause selbst ausfüllen	2.8%

Tabelle 3: Eingeständenes Täuschungsverhalten ²³

¹⁹ siehe auch in: HERRY M.: EDV-Software - ein hilfreiches Instrument für die Verkehrsplanung? Vortrag zum Symposium CORP'96, Tagungsband, S.205-218, Wien 1996

²⁰ siehe zum Beispiel in: SACHS L.: Angewandte Statistik. Springer-Verlag, 1992

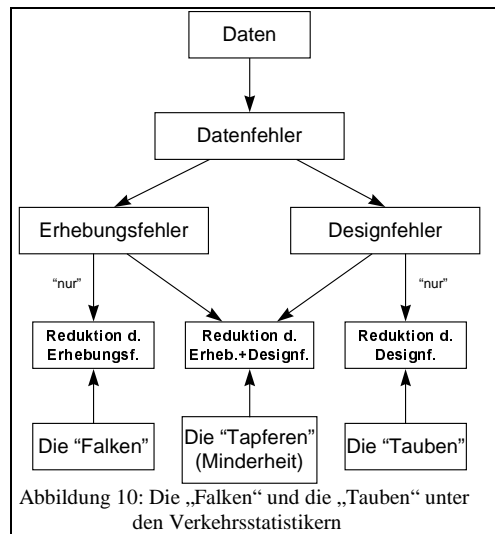
²¹ siehe zum Beispiel in: KRUG W., NOURNEY M.: Wirtschafts- und Sozialstatistik - Gewinnung von Daten. R.Oldenbourg Verlag München Wien, 1982

²² KÖLTRINGER R.: Gültigkeit von Umfragedaten. Böhlau Verlag Wien - Köln - Weimar, 1993

Leider muß festgestellt werden, daß es - zumindestens in der empirischen Verkehrsforschung - keine einheitliche Fachmeinung über die einzuschlagenden Wege zur Reduktion von Datenfehlern gibt. Vielmehr kristallisieren sich **zwei Richtungen** heraus:

- die „Falken“ unter den Verkehrsstatistikern, die sich auf die Bereinigung der Zufallsfehler mit den formalen Methoden der Theorie der Standardfehler konzentrieren und „vor lauter Standardfehler“ die Designfehler aus den Augen verlieren - sich auch gern als die Gralshüter der „reinen“ Statistik sehen, sowie
- die „Tauben“ - die „Sozialarbeiter“ unter den Statistikern, die sich vornehmlich mit den Design-Fehlern auseinandersetzen, von formaler (mathematischen) Statistik aber wenig halten.

Die Menge von Verkehrsforschern und -planern, die sich intensiv mit beiden Bereichen, nämlich den Zufalls- und den Design-Fehlern, beschäftigen und ihr Wissen dann auch in der Praxis umsetzen können, ist hingegen sehr klein.



3.2.2. Datengenauigkeit von analytischen Daten

Die Datengenauigkeit von analytischen Daten beinhaltet die **Fehlerfortpflanzung** in den die analytischen Daten erzeugenden Modellen, zum Beispiel den Verkehrsmodellen.

Dies kann auf zweierlei Art und Weise realisiert werden:

- erstens, durch Aufstellung der **modellspezifischen Fehlerfortpflanzungs-Funktion** und
- zweitens, mittels **Sensitivitätsanalysen**.

Zu den Fehlerfortpflanzungs-Funktionen ist folgendes anzumerken:

- Die Ermittlung solcher Funktionen ist nicht einfach,
 - weil sie nicht nur gediegene Kenntnis über die Fehlerfortpflanzungstheorie erfordert, derer sich anzueignen es fast kein Verkehrsplaner oder -forscher der Mühe wert empfindet,
 - weil sie auch die konkrete Kenntnis der Modellalgorithmen und -operationen voraussetzt, was jedoch nur selten der Fall ist, da
 - * die verwendeten Modelle entweder für den Benutzer „Black-Boxen“ darstellen oder
 - * die Modelle zwar mathematisch aber nicht verständlich „lesbar“ sind.
- Hat man es der Mühe wert empfunden und geschafft, die Fehlerfortpflanzungs-Funktionen zu erhalten, ist erst die „Hälfte“ des Weges auf der Fehlersuche von analytischen Daten zurückgelegt. Die zweite Hälfte besteht in der „richtigen“ Anwendung dieser Funktion und der sachgemäßen Interpretation ihrer Ergebnisse.

3.2.3. Wie genau ist „genau genug“?

Oft wird die allzu berechtigte Frage gestellt, wie genau Daten für (wichtige) Untersuchungen sein müssen, das heißt, was ist „genau genug“?

Wir wollen dieser Frage im Rahmen unserer Themenstellung nachgehen. Ausgangspunkt ist dabei das INFRAS-Essay-Papier zu diesem Thema für die sozialen Kosten und Nutzen im Verkehrsbereich ²³.

Das Kernproblem ist bei vielen Fragestellungen, daß „zu hohe“ Anforderungen an die Datengenauigkeit die Gesamtaussage in Frage stellen! Das mag zunächst widersprüchlich klingen, ist aber mitunter so: Nehmen wir als Beispiel die Berechnung von Wegekostendeckungsgraden im Verkehrswesen ²⁴. Dazu stellt INFRAS fest: „Je ‘wissenschaftlich exakter’ die Zahlen fundiert sein ‘müssen’, desto größer ist die Gefahr, daß die eigentliche ‘Kostenwahrheit’ unterschätzt wird, weil wichtige - aber schlecht quantifizierbare - Bereiche weggelassen werden ‘müssen’.“

Das hat also damit zu tun, daß sich Genauigkeiten von „Gesamt“-Aussagen

- nicht nur aus den Genauigkeiten der Teilbereiche zusammensetzen, die zu der Gesamtaussage führen,
- sondern auch daraus ergeben, welche Information dazu weggelassen wurden/werden mußten.

Der erste Bereich hängt eng mit der **Fehlerfortpflanzung** - siehe Kapitel 3.2.2 - zusammen, der zweite mit dem Komplex „Arbeiten mit unvollständigen oder fehlenden Informationen“, auf das im Kapitel 3.3 näher eingegangen wird.

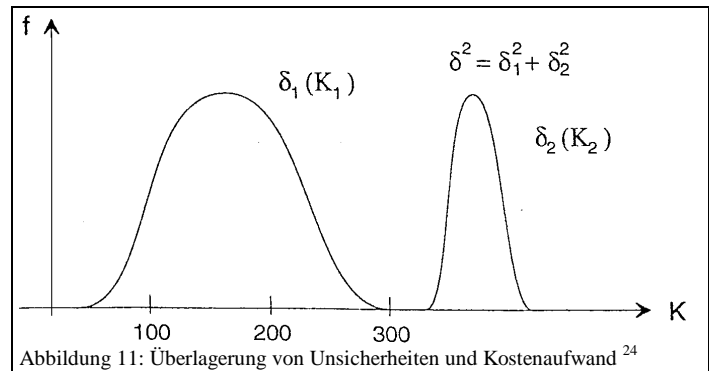
Schließlich hat dieses Problem aber auch etwas mit dem Kosten- und Zeitaufwand zu tun: Je genauer schlecht quantifizierbare Größen ermittelt werden sollen, desto teurer und zeitraubender (und unter Umständen unergiebig) werden die dazu benötigten Forschungsarbeiten.

²³ INFRAS: Wie gau ist genau genug? Soziale Kosten und Nutzen im Verkehrsbereich. Essay-Papier im Auftrag des Dienstes für Gesamtverkehrsfragen der Schweiz, Zürich, 1993

²⁴ siehe zum Beispiel in: HERRY M.: Wegekosten und Maut in Österreich. In: Schriftenreihe der Institute für Eisenbahnwesen sowie Straßenbau und Verkehrswesen der TU Graz, Heft 15, Graz 1993

Folgende Ableitungen können getroffen werden:

- Vom Standpunkt der statistischen Genauigkeit ist die Frage nach dem „genau genug“ mit der Höhe der Sicherheitswahrscheinlichkeit von Aussagen und der Fehlerfortpflanzung (das heißt auch Fehlerüberlappung) zu beantworten.
- Andererseits ist diese Frage auch zu relativieren: Wie genau genug **für welche** (fachlichen oder verkehrspolitischen) **Entscheidungen!** Die Genauigkeit muß mindestens so hoch sein, daß
 - das Risiko der daraus abzuleitenden Entscheidungen
 - * fachlich,
 - * verkehrspolitisch und
 - * gesellschaftlich akzeptabel ist,
 - das Risiko der Aufschiebung von daraus abzuleitenden Entscheidungen aufgrund einer genaueren Erforschung der Ergebnisse zu hoch ist.
- Das kann man auch so ausdrücken: Mit welchen (begrenzten) gegebenen Aufwand gelangt man zu einer möglichst hohen Aussagekraft der Ergebnisse, sprich, zu einem möglichst hohem Informationsgehalt.
- Das „genau genug“ hängt schließlich sehr stark
 - mit dem Differenzierungsgrad der Aussagen und
 - der Methode, wie man zu der Gesamtaussage kommt:
 - * deduktiv oder
 - * induktiv,
 zusammen.
 Nach dem Motto: Abgeleitete Aussagen sind weniger aufwendig (und problematisch) als „aufsummierte“.



3.3. Verarbeitung von unvollständigen oder fehlenden Informationen

Die Verarbeitung von unvollständigen oder fehlenden Informationen ist insofern ein äußerst wichtiges Kapitel im Zusammenhang mit der Betrachtung von Datenqualität als

- in vielen Bereichen - insbesondere aber bei der Gewinnung von empirischen Daten - die zur Verfügung stehenden Informationen
 - unvollständig, die sogenannten „k.A.“ ‘s, das heißt „keine Angaben“, sind oder
 - überhaupt fehlen,
- diesem Umstand jedoch zu wenig Aufmerksamkeit gewidmet wird, was in der Regel zu fatalen Auswirkungen führt.

3.3.1. Arbeiten mit unvollständigen Informationen

Unvollständige Informationen können:

- gewollt „herbeigeführt“ werden (!) oder
- faktisch „ungewollt“ im Verlaufe der Datenerhebung entstehen.

Der erste Aspekt ist eine **aktive Erzeugung von unvollständigen Informationen** und hängt mit der im Kapitel 3.2.3 untersuchten Fragestellung „Wie genau ist ‘genau genug’?“ zusammen.

Der zweite Aspekt ist eine **passive Erzeugung von unvollständigen Informationen** und hat etwas mit den Fehlern in empirischen Daten zu tun, der allgemein bereits im Kapitel 3.2.1.2 behandelt wurde, und im besonderen im nächsten Kapitel 0 ausgeführt wird.

Zur Behandlung von „keinen Angaben“ ist folgendes **prinzipiell** zu sagen:

- Die „keine Angaben“-Fälle **müssen** in den Auswertungen und Ergebnissen **immer**
 - berücksichtigt und
 - auch aufgeführt werden, was leider nur selten praktiziert wird.
Das ist deshalb so wichtig, da sonst die (Auswertungs-)Grundgesamtheiten **nicht gleich** sind, was unweigerlich zu **falschen** Interpretationen und Schlußfolgerungen führt (siehe Kapitel 3.4.4).
- Die „keine Angaben“-Anteile bei den Werten einer Variablen sollten die 5%-Marke nicht überschreiten. Andernfalls
 - kann dies ein Hinweis auf Design-Fehler in der Datenerhebung sein (vergl. mit Kapitel 3.2.1.2) bzw.

- können sie zu Interpretations- und Weiterverarbeitungsschwierigkeiten führen.
- „Keine Angaben“-Fälle werden oft mit „Trifft nicht zu“-Angaben verwechselt:
 - „Keine Angabe“ liegt vor, wenn eine Angabe gemacht werden hätte müssen, sie aber nicht gemacht wurde oder nicht vorliegt (aus welchen Gründen auch immer),
 - während eine „Trifft nicht zu“-Angabe bedeutet, daß von (bestimmten) Teilen der Erhebungsgesamtheit dazu keine Angaben gemacht werden können.

Zum Beispiel kann ein 6-jähriger Bub in Österreich keinen Führerschein haben, das heißt, wenn er zur Frage „Besitzen Sie einen Führerschein?“ nicht antwortet, so ist das nicht als „Keine Angabe“ zu behandeln, sondern als „Trifft nicht zu“-Fall!

Eine Möglichkeit, vorliegende „keine Angaben“ zu korrigieren, besteht in dem Ersetzen dieser nicht bestehenden Angaben durch die vorliegenden:

- Das kann
 - entweder durch Mittelwertübertragung oder
 - durch „intelligente“ Einzelwertübertragung realisiert werden.
- Dabei wird jedoch allzu oft übersehen, daß die erste Methode mittelwertsinvariant ist, die zweite aber die betreffenden Mittelwerte in der Regel nicht gleich hält!

Eine radikale Methode bei der Auswertung von Daten mit „keine Angaben“ wird mitunter derart angewandt, daß sämtliche Fälle, bei denen in irgendeiner Variablen „k.A.“-Werte vorkommen, herausgenommen werden. Dies ist jed. sehr gefährlich, da

- erstens, die Stichprobe in der Regel wesentlich reduziert wird und
- zweitens, die Stichprobenstruktur und der Stichprobeninhalt dadurch völlig zerstört werden können.

Zum Abschluß der Hinweise für das Arbeiten mit unvollständigen Information bringen wir noch einige Bemerkungen zu einem speziellen Bereich der unvollständigen Angaben im Verkehrsbereich, den „non-reported trips“.

3.3.2. Non-Reported-Trips-Analyse

Die Non-Reported-Trips-Analyse beschäftigt sich mit unvollständiger Information, und zwar mit nicht berichteten Wegen. Sie spielt damit vor allem in der Mobilitätsforschung eine wichtige Rolle, die jedoch oft unterschätzt und zu wenig Beachtung bei den Mobilitätserhebungen findet. Immerhin kann der Anteil der nicht berichteten Wege bis zu 30% betragen!

Nicht berichtete Wege beziehen sich vor allem auf

- die Nach-Hause-Wege und
- kurze Wege.

Die folgenden Auswertungen sind der Arbeit von BRÖG / ERL / MEYBURG ²⁵ entnommen und bringen einen Überblick über die Verteilung der nicht berichteten Wege, differenziert nach

- Wegelänge und -dauer,
- benutztem Verkehrsmittel und
- Wegezweck.

Trip Length (in km)	Reported Information (in %)	Explored Information (in %)	Non-Reported Trip Rate (in %)
0- 0.4	11.0	3.9	26.3
0.5-0.9	9.4	2.9	23.4
1.0-2.9	23.9	3.8	13.9
3.0 -4.9	10.7	1.2	9.9
5.0 -9.9	16.6	1.4	7.8
10.0 -19.9	7.1	.6	7.7
above20.0	7.1	.4	5.3
TOTAL	85.8	14.2	(14.2)
Average Trip Distance	5.7 km	3.3 km	8.8%**
Average Trip Duration	20.4 mins.	15.9 mins.	11.4%***
** = $\frac{\text{(Number of "explored" trips)} \times \text{(Average length for "explored" trips)}}{\text{(Total number of trips reported and explored)}} \times \text{(Average trip length)}$			
*** = [Computation equivalent to **]			

Abbildung 12: Anteil der nicht berichteten Wege, differenziert nach Wegelänge und -dauer

Trip Purpose	Reported Information (in %)	Explored Information (in %)	Non-Reported Trip Rate (in %)
Work	19.0	1.2	5.9
School (Training)	6.7	.6	8.2
Shopping	22.7	5.1	18.4
Other Discretionary Activities	11.8	2.0	14.5
Recreation	25.6	5.3	17.2
TOTAL	85.8	14.2	(14.2)
Purpose Aggregation			
Regular Activities	25.7	1.8	6.4
Discretionary Activ.-Recreational	34.5	7.1	17.1
Activ.	25.6	5.3	17.2

Abbildung 14: Anteil der nicht berichteten Wege, differenziert nach Wegezweck

Predominantly Used Travel Mode	Reported Information (in %)	Explored Information (in %)	Non-Reported Trip Rate (in %)
Walking	26.4	7.9	22.9
Bicycle	7.8	1.3	14.4
Moped, Motorcycle, Motorcycle	1.2	.4	25.0
Auto Driver	30.3	2.9	8.9
Auto Passenger	6.1	.8	12.3
Public Transit	12.6	-9	6.7
Train	1.4	-	0.0
TOTAL	85.8	14.2	14.2
Mode Aggregations			
Non-Motorized Modes	35.2	9.2	21.1
Motorized Modes	37.6	4.1	9.9
Public Transport	14.0	.9	6.1

Abbildung 13: Anteil der nicht berichteten Wege, differenziert nach benutztem Verkehrsmittel

Activities	Non-Motorized Travel			Motorized Travel			Public Transit		
	RI*	EI**	NRTR***	RI	EI	NRTR	RI	EI	NRTR
Regular									
Trip Frequency	7.6	1.0	11.5	11.0	.4	3.4	7.1	.4	5.3
Trip Length (km)	1.4	.7	6.1	10.7	7.4	2.4	9.0	9.2	5.4
Trip Duration (min)	13.3	8.1	7.4	22.5	15.0	2.3	37.6	40.0	5.7
Discretionary									
Trip Frequency	17.2	4.8	21.8	14.2	2.2	13.6	3.2	0.1	3.9
Trip Length (km)	.9	.6	16.2	6.8	4.6	9.6	9.2	10.0	4.2
Trip Duration (min)	11.6	8.3	11.7	15.6	14.5	12.6	44.6	60.0	4.0
Recreational									
Trip Frequency	9.4	3.4	26.5	12.4	1.5	10.8	3.7	.4	9.5
Trip Length (km)	1.1	.9	17.4	8.4	7.8	9.3	9.8	8.0	5.1
Trip Duration (min)	15.0	20.9	32.6	24.1	25.5	11.3	41.9	33.8	8.0

RI = Reported Information; EI = Explored Information.; NRTR = Non-reported trip rate

Abbildung 15: Anteil der nicht berichteten Wege, differenziert nach nach benutztem Verkehrsmittel und Wegezweck

3.3.3. Arbeiten mit fehlenden Informationen

Fehlende Information liegen (analog zu den unvollständigen Informationen)

- entweder bewußt
- oder unbewußt

vor.

Der **bewußte** Fall tritt in dem großen Bereich der Stichprobenbildung ein: Man will (oder kann) gar nicht alle Information (Grundgesamtheit) erhalten, sondern „begnügt“ sich mit Daten von einer Teilmenge, eben der Stichprobe, in der nicht alle Elemente der Grundgesamtheit enthalten sind.

Zu beachten ist dabei

- nicht nur der Aspekt der Teilmenge (der Grundgesamtheit), die nicht in der Stichprobe liegt,
- sondern auch jene Ausfälle **innerhalb** der Stichprobe, wie
 - die unechten Ausfälle (Auswahl-Stichprobe --> Brutto-Stichprobe),
 - die echten Ausfälle (Brutto-Stichprobe --> Antworter-Stichprobe) und
 - die nicht verwertbaren Fälle (Brutto-Stichprobe --> Netto-Stichprobe).

Beide Aspekte sind für die Hochrechnung (auf die Grundgesamtheit) sehr wichtig.

Der Aspekt der Stichprobenbildung ist in dem Vortrag von HERRY zum CORP'96 dargestellt ²⁶.

3.3.4. Non-Response-Analyse

Die Non-Response-Analyse beschäftigt sich mit den echten Ausfällen (Brutto-Stichprobe --> Antworter-Stichprobe), das heißt, versucht die fehlenden Informationen der Nichtantworter (der Brutto-Stichprobe!) ²⁷ bei den Auswertungen mit zu berücksichtigen, und zwar in der Hochrechnung der (Erhebungs-)Ergebnisse von der Netto-Stichprobe auf die Brutto-Stichprobe ²⁷.

Dieser Aspekt spielt nicht nur, aber vor allem in schriftlich-postalischen Erhebungen eine bedeutende Rolle.

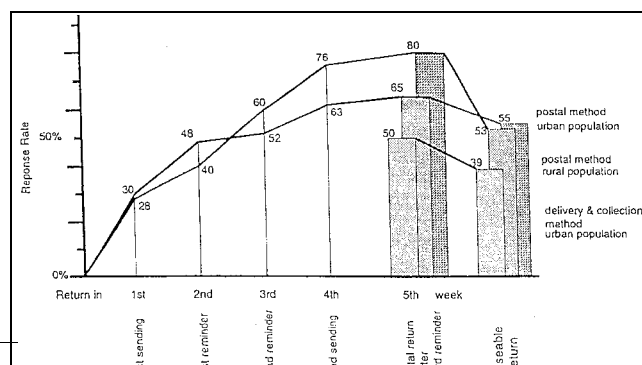
Trotzdem muß erwähnt werden, daß in vielen Untersuchungen dieser Aspekt entweder nur ungenügend oder sogar überhaupt nicht berücksichtigt wird:

- In diesen Fällen wird nämlich von der Netto-Stichprobe auf die Brutto-Stichprobe **proportional hochgerechnet**,
- was in vielen Fällen zu **falschen** Ergebnissen führt,
- weil sich die Nichtantworter - vor allem im Verkehrsbereich - eben nicht so verhalten wie die Antworter.

Die folgende Abbildung zeigt Kurven zum Rücklaufverhalten in Abhängigkeit

- von der Zeit / der Erhebungsaktion und
- verschiedenen Befragungsmethoden

für eine Erhebung zur täglichen Mobilität.



²⁵ BRÖG W., ERL E., MEYBURG A.H.: Problems of Non-reported Trips Transportation Research Board, Washington, D.C., January, 18-22, 1982

²⁶ HERRY M.: EDV-Software - ein hilfreiches Instrument für die Verkehrsplanung - Vortrag zum Symposium CORP'96, S. 205-218, Kapitel 2, Wien 1996

²⁷ Es gibt auch noch andere Nichtantworter!

²⁸ SAMMER G., FALLAST K., KÖSTENBERGER H. et al.: Verkehrsuntersuchung zum Nahverkehrskonzept Salzburg 1982. Im Auftrag der Salzburger Landesregierung et al., Graz 1985

Für eine Fernverkehrserhebung (Pilotstudie) ist die Rücklaufkurve in der Abbildung 17 dargestellt.

Non-Response-Funktionen zu einem Verkehrsmerkmal V zeigen die Abhängigkeit der Variablen V vom (zeitlichen) Rücklaufverhalten. Hieraus ist deutlich sichtbar,

- wie unterschiedlich das Verkehrsverhalten von Nichtantwortern im Vergleich zu den Antwortern ist,
- daß die Non-Response-Funktionen für ein und dasselbe Verkehrsmerkmal, differenziert nach verschiedenen Verkehrsegmenten, durchaus sehr unterschiedlich sein kann.

POLAK / AMPT³⁰ bringen neuere Untersuchungsergebnisse aus englischen und australischen Verkehrsbefragungen zu diesem Thema.

Einen interessanten - wenn auch nach meiner Meinung nicht unumstrittenen - Weg zur Non-Response-Analyse wird bei ARMOOGUM / MADRE³³ besprochen: Sie filtern bezüglich verschiedener soziodemografischer Merkmale Personengruppen heraus, die ein gleiches Rücklaufverhalten (bezüglich des Verkehrsmerkmals V) aufweisen und rechnen dann für die Personenmengen das Merkmal V proportional von der Netto- auf die Brutto-Stichprobe hoch.

3.4. Daten(-Interpretation) als Information

Ein weiterer wichtiger Zusammenhangsbereich zur Beziehung Daten - Information ist der Aspekt „Daten als Information“. Wir wollen diesen Bereich unter folgenden Aspekten kurz beleuchten:

- Informationsgehalt - Entropie,
- Perzeption und Wirklichkeit,
- Datenmengen und Information und
- Gefahren bei der Datendarstellung (Wie lügt man mit Statistik).

3.4.1. Informationsgehalt - Entropie

In der Physik wird das Maß der „thermodynamischen Wahrscheinlichkeit“ eines Zustands als Entropie bezeichnet.

Die thermodynamische Wahrscheinlichkeit ist ein Maß der Unordnung; in einem sich selbst überlassenen geschlossenen System nimmt sie und damit die Unordnung ständig zu. Die thermodynamische Wahrscheinlichkeit W ist nun mit der Entropie H des Systems nach BOLZMANN durch die Beziehung verknüpft:

$$H = k * \ln W \quad (F3)$$

Das besagt: Die Entropie ist der thermodynamischen Wahrscheinlichkeit proportional (k, der Proportionalitätsfaktor, ist die BOLZMANNsche Konstante), das heißt, ist W groß, so ist auch die Entropie des Zustandes groß; die Entropie wächst mit der Wahrscheinlichkeit des Zustandes.

Das ist nun aber genau umgekehrt wie bei der Information: je wahrscheinlicher der Zustand, desto geringer ist die Information, die er enthält, und je unwahrscheinlicher, desto größer sein Informationsgehalt.

Es ist ein allgemeines, übrigens auch aus der alltäglichen Erfahrung bekanntes, Prinzip der Natur:

- Unordnung ist wahrscheinlicher als Ordnung.
- Die Unordnung nimmt von allein zu:
 - Das gilt nun, worauf SHANNON und WIENER besonderen Wert legen, auch für die Übertragung von Signalen,
 - und das äußert sich darin, daß „eine Nachricht ihre Ordnung während des Aktes der Übertragung wohl von selbst verlieren, aber nicht gewinnen kann“³⁴.

²⁹ AXHAUSEN K., KÖLL H., HERRY M.: LongDistance Travle Demand Measurement Methods Pilot Study. Im Auftrag des Bundesministeriums für Wissenschaft, Verkehr und Kunst, Innsbruck, Wien 1996

³⁰ POLAK J.W., AMPT E.S.: An Analysis of Response Wave and Nonresponse Effects in Travel Diary Surveys. Paper presented at the 4th International Conference on Survey Methods in Transport, Steeple Aston, 9-11 September 1996

³¹ HERRY M.: Gewichtung der KONTIV 1982. In: Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft, B85, 1986

³² HERRY M.: Auswirkungen und Behandlung von Datenschutzproblemen bei haushaltsbefragungen zum Verkehrsverhalten. In: Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft, B87, 1987

³³ ARMOOGUM J., MADRE J.-L.: Nonresponse Correction in the French 1993-94 NPTS: The Example of Daily Trips. INRETS, Paris, 1996

³⁴ WIENER N.: Cybernetics. Paris, New York 1948

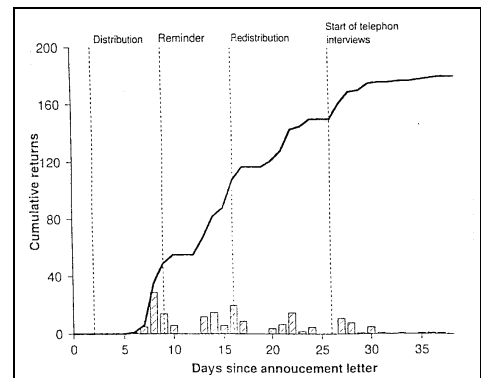


Abbildung 17: Fragebogenrücklauf in Abhängigkeit von der Zeit / der Erhebungsaktion für eine Fernverkehrsbefragung²⁹

Non-Response-Kurven-Typen für Wege Pro mobiler Person

	REGEL- MÄSSIGER VERKEHR	GELEGEN- HEITS- VERKEHR	FREIZEIT- VERKEHR	SUMME
NMV	—	↘	↘	↘
MIV	∪	∩	∩	∩
ÖV	—	—	—	—
SUMME	—	∩	↘	∩

— (annähernd) konstant; ↘ (annähernd) monoton fallend

∩ (annähernd) konkav gewölbt; ∪ (annähernd) konvex gewölbt

Abbildung 18: Non-Response-Funktionen zum Verkehrsmerkmal „Wege pro mobiler Person“ in der KONTIV'82^{31 32}

- Da Information nun auf Ordnung beruht, so kann ein Maß der Ordnung - etwa einer Zeichenfolge - zu einem Maß der Information werden.
- Andererseits ist die Entropie ein Maß der Unordnung bzw. einer stets zunehmenden Unordnung.
- Das Negative der Entropie, die sog. Negentropie, kann dann als ein Maß der Ordnung, und zwar einer ständig abnehmenden Ordnung angesehen werden.

Analog verhält es sich mit einer Informationen: Zum Ereignisfeld $E = (E_1, \dots, E_n)$ mit den Ereignissen E_i und ihren Eintrittswahrscheinlichkeiten p_i lautet die durchschnittliche Informationsmenge H (Informationsgehalt) ³⁵

$$H = - \sum p_i \log p_i \quad (F4)$$

Die Entropie hat in der Verkehrsforschung eher nur eine geringe Spur hinterlassen, was eigentlich verwunderlich ist, spielt doch die Informationsmenge in den „Verkehrsdingen“ - wie wir bisher bereits sehen konnten - eine nicht unerhebliche Rolle.

Einen stärkeren Niederschlag findet dieser Begriff in den Entropie-Verkehrsmodellen ³⁶, die jedoch in der Verkehrsplanung eine untergeordnete Rolle spielen.

Wir wollen nun des weiteren diesen Bereich unter den Aspekten der Informationsperzeption, der Datenmenge und der Datendarstellung kurz betrachten.

3.4.2. Perzeption und Wirklichkeit

Die Informationen, die wir als Abbildung der uns umgebenden Wirklichkeit empfangen, haben keinen absoluten, sondern mitunter einen sehr relativen, subjektiven Charakter.

So wird heute (noch) vielfach der öffentliche Verkehr als langsam, teuer und primitiv empfunden, während das Auto als schnell, billig und komfortabel gilt.

Es klafft also zwischen der Perzeption der Wirklichkeit und der „wirklichen“ Wirklichkeit eine zum Teil empfindliche Lücke, wobei man fragen kann, wie wirklich die Wirklichkeit wirklich ist ³⁶, das ist so etwas, wie die „optischen“ Täuschungen:

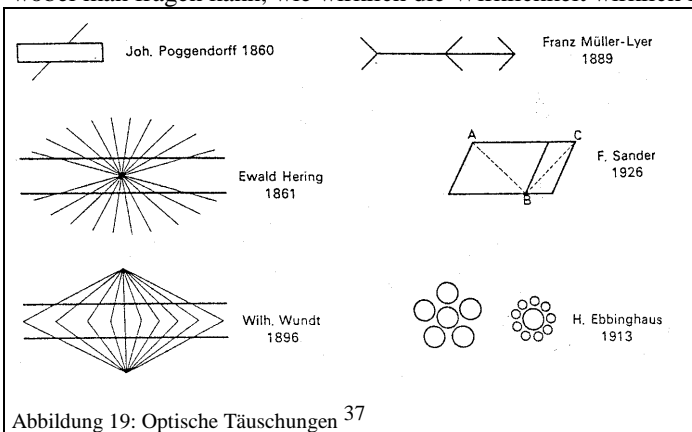


Abbildung 19: Optische Täuschungen ³⁷

Was nun das Verkehrsverhalten anbelangt, so muß man zwischen

- Einstellung (zum Verkehr) und
- tatsächliches Verhalten (im Verkehr)

unterscheiden.

Die folgende Abbildung 20 zeigt recht deutlich die Unterschiede zwischen diesen beiden Kategorien.

Daraus folgt nun bei Leibe nicht,

- daß für uns die Einstellungen zum Verkehr uninteressant sind,
- im Gegenteil, weil die Verkehrsentscheidungen der BürgerInnen eben durch diese Einstellungen getroffen werden,
- das heißt,
 - wir müssen nicht nur das tatsächliche Verhalten kennen,

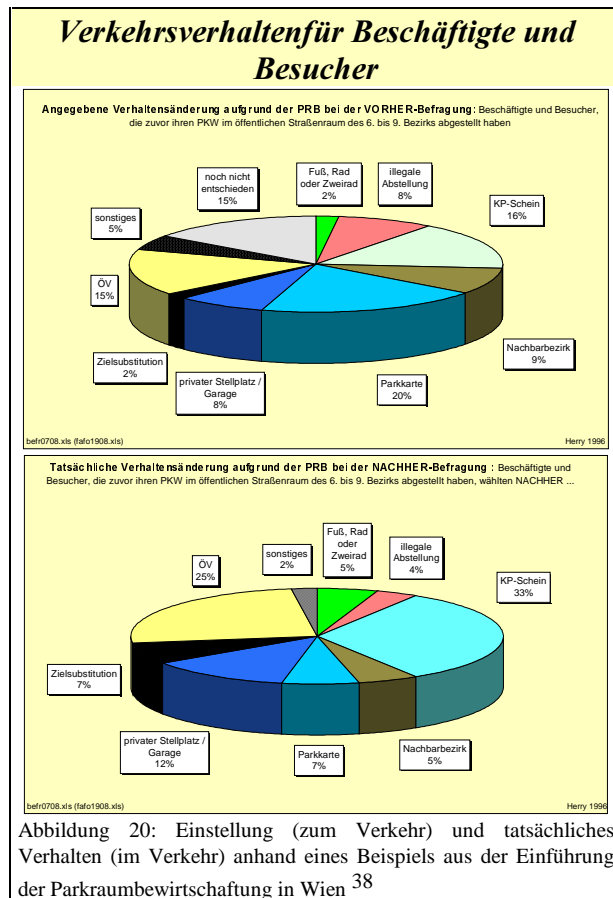


Abbildung 20: Einstellung (zum Verkehr) und tatsächliches Verhalten (im Verkehr) anhand eines Beispiels aus der Einführung der Parkraumbewirtschaftung in Wien ³⁸

³⁵ siehe zum Beispiel: RICHTER K.-J.: Der Entropieansatz in der Verkehrsanalyse und in der Verkehrsplanung. In: Österreichische Zeitschrift für Verkehrswissenschaft, Heft 1/90, S. 23-29

³⁶ WATZLAWICK P.: Wie wirklich ist die Wirklichkeit. Piper-Verlag München - Zürich, 1978

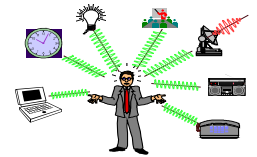
³⁷ aus: BAUER F.L., GOOS G.: Informatik. Springer-Verlag Berlin-heidelberg-New York, 1982

³⁸ HERRY M., ROSINAK W.: Vorher-Nachher-Untersuchung zur Einführung der Parkraumbewirtschaftung in den Bezirken 6 bis 9. Im Auftrag des Magistrats der Stadt Wien, Wien 1997

- sondern auch den dazugehörigen (verkehrspsychologischen) Hintergrund!

3.4.3. Datenmengen und Information

Oft wird - und das insbesondere im Verkehrsbereich - über akuten Mangel an (Verkehrs)-Daten geklagt. Nun, ich würde sagen, gar nicht mal so selten ist das umgekehrt: Wir haben „zu viele“ Daten - in dem Sinne, daß ein und dieselben verkehrlichen Sachverhalte durch mehrere Datenmengen abgebildet werden. Im Prinzip ist das nicht schlecht, weil man dann - insbesondere bei „heiklen“ Daten - Vergleiche hat und den Datenbestand dann entsprechend evaluieren kann.



Aber genau darin liegt auch das Problem. Mehrere Daten zu einem verkehrlichen Aspekt beinhalten auch Risiken:

- Zu viele Daten - Datenfriedhöfe - lassen den „Wald vor Bäumen“ nicht sehen, sie verstellen den Blick für das Wichtige und Wesentliche (**Datendichte-Streß**).
- Bei mehreren Daten zu ein und demselben Sachverhalt (vergl. zum Beispiel **Fehler! Verweisquelle konnte nicht gefunden werden.**) treten fast immer Unterschiede auf, die dann
 - bewertet und
 - eine anschließende Datenharmonisierung verlangen, die oft schwierig, zeitraubend und kostenintensiv ist und letzten Endes zu einem weiteren Datenbestand führt!
- Auch unter zeit- und kostenökonomischen Gründen ist diese Situation unbefriedigend.
- Dazu kommt noch, daß in der Regel Eichungsvorgänge bei dem einen Datenbestand aus Informationen des anderen Bestands bestehen, die Datenmenge zwar groß ist, der Informationsgehalt aber relativ bescheiden ist.

Wichtig erscheint mir in diesem Zusammenhang, relativ „zentral“ zu klären,

- **welche Information/Daten**
 - wirklich und
 - zu welchem Zweck **gebraucht** werden,
- **wie genau** und umfangreich diese Information/Daten sein müssen und
- **wer sie**
 - **produzieren** und
 - **erhalten** soll.

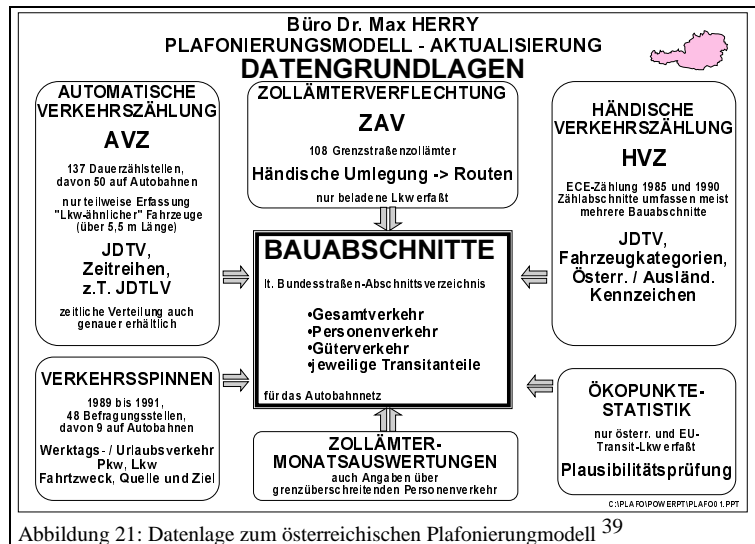


Abbildung 21: Datenlage zum österreichischen Plafonierungsmodell ³⁹

3.4.4. Wie lügt man mit Statistik

Ganz wesentlich für die Beziehung „Daten als Information“ scheint mir ihre **Darstellung**. Die folgenden beiden (schematischen) Abbildungen zeigen, was ich darunter verstehe.

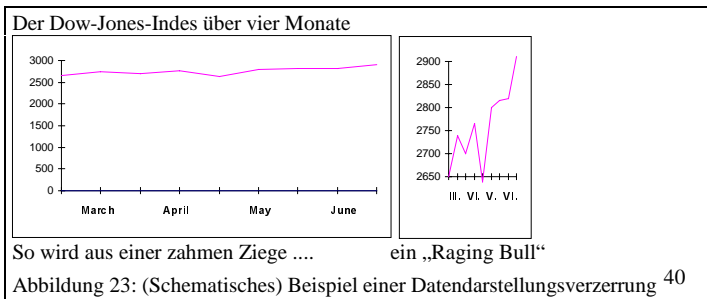
Unabhängig, ob gewollt (= Manipulation) oder nicht gewollt (= Einfältigkeit), sollten wir durch die Darstellung von Daten und Informationen nicht noch falsche Informationen hinzufügen.

Noch zwei weitere Beispiele aus der Verkehrspraxis:

- Abbildung 24 zeigt eine offizielle Statistik der EU. Darin sind die Modal-Split-Werte für europäische Länder ausgewiesen. Am österreichischen Wert fällt der extrem hohe Anteil des Schienenverkehrs auf - der, in der Tat, wirklich, zumindestens in diesem Ländervergleich, sehr gut ist, aber eben nicht so hoch - er ist nämlich falsch:
 - Ursache dafür ist, daß für den Schienengüterverkehr der Gesamtverkehr genommen wurde, für den Straßengüterverkehr jedoch nur der Fernverkehr!
 - Aber auch das ist von der „Genesis“ her erklärbar: Der Straßengüter-Nahverkehr wird nur alle 5 Jahre erhoben, das Tabellen-Jahr war aber nun ein Jahr, für das kein Nahverkehr erhoben wurde, der (erhobene) „Gesamtverkehr“ also wirklich nur aus dem Fernverkehr bestand.

³⁹ HERRY M.: Österreichisches Verkehrs-Plafonierungsmodell. Im Auftrag des Bundesministeriums für öffentliche Wirtschaft und Verkehr, Wien 1995

- Der Umstand, der zu dieser fatalen Situation führte, ist also ein höchst „formaler“ Fehler, der sich jedoch peinlich auswirkt.

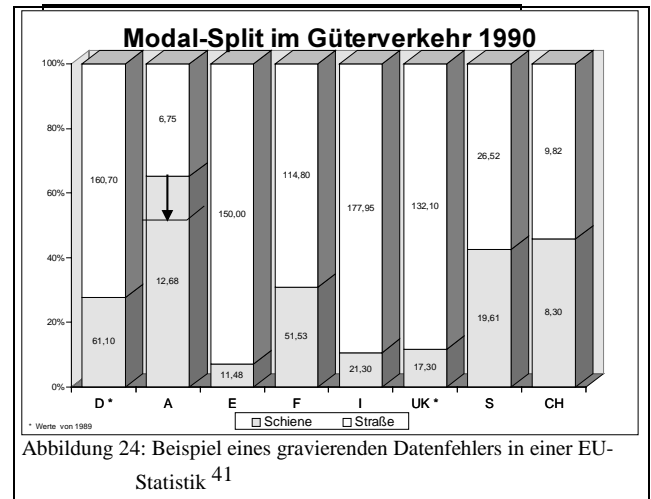


- Im Straßengüterverkehr galt in den letzten Jahren folgende Faustregel: Der Anteil des Transitverkehrs beträgt
 - bezüglich der Fahrleistung bei ca. 5%,
 - bezüglich der Tonnage bei ca. 15% und
 - bezüglich der Transportleistung bei ca. 25%⁴²

Diese sehr unterschiedlichen Zahlen, die allein durch die entsprechende Dimension geprägt sind, werden nun in jeder Variante ohne Nennung der Dimension kolportiert:

- Der Transit-„Gegner“ sagt, der Transitverkehr durch Österreich betrage 25%,
- der „Befürworter“ meint, nein, er ist ja nur 5%.
- Das ist aber das 5-fache!

Informationen bestehen also nicht nur aus „nackten“ Zahlen, sondern sind mit einem wichtigen Kontext versehen, die diese Daten erst „wahr“ (oder „falsch“) machen!



3.5. Informationen über Daten - Informationszugang

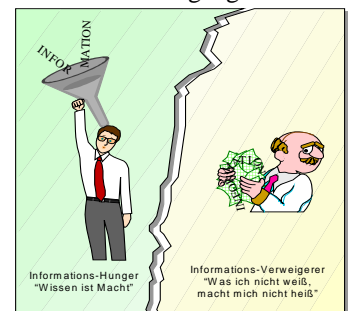
Der Bereich „Informationen zu den oder über die Daten“ ist ein relativ schwieriger, da er den Informationszugang betrifft. Wir wollen ihn kurz unter folgenden Gesichtspunkten behandeln:

- Informations-„Hunger“,
- Informationsangst und
- Informationsschutz.

3.5.1. Informationslust - „Wissen ist Macht“

(Fast) alle wissen, wie wertvoll und wichtig es ist, „richtige“ Informationen

- zu besitzen und
- sie auch weitergeben dürfen.



Nicht von ungefähr halten nichtdemokratische Staaten viele Information geheim bzw. manipulieren sie. Daten nicht weiterzugeben,

- muß jedoch nicht nur aus „übergeordnetem“ Interesse entspringen,
- sondern auch aus privatem oder (privat-)wirtschaftlichen.

Ich kann mich noch gut erinnern, vor 15 Jahren war es nicht üblich, unter Kollegen Daten auszutauschen (vielleicht gab es damals keine (Verkehrs-)Kollegen). Heute ist das schon anders, aber eben nur unter Kollegen!

Ich würde sehr dafür plädieren, daß alle (nichtdatenschutzbedürftigen) Informationen - insbesondere die, welche mit Steuergeldern finanziert wurden - an alle Institutionen und Personen weitergegeben werden sollten - unter Entrichtung einer Schutzgebühr, um sinnlose Entnahmen zu verhindern.

Oft wird dazu der Vorschlag gemacht: nicht an alle, sondern nur an die „Kompetenten“.

Nur:

- Erstens, wer ist (Daten-)kompetent und
- zweitens, wer entscheidet, wer kompetent ist?

Das löst das Problem sicher nicht.

Ein weiterer Punkt in diesem Sinn ist die Tatsache, daß zwar viele öffentliche Ämter und Verwaltungen willens sind, „ihre“ Daten weiterzugeben, das aber nicht können, weil sie nicht über diese Daten verfügen, sie haben sie physisch einfach nicht!

⁴⁰ KRÄMER W.: So lügt man mit Statistik. Campus-Verlag Frankfurt - New York, 1995

⁴¹ HERRY M., ROSINAK W.: Vorher-Nachher-Untersuchung zur Einführung der Parkraumbewirtschaftung in den Bezirken 6 bis 9. Im Auftrag des Magistrats der Stadt Wien, Wien 1997

⁴² HERRY M.: Der Güterverkehr in Österreich. ÖROK-Schriftenreihe, Heft 100, Wien 1994

Das mag verwundern, aber bei der Mehrzahl der Datenbeauftragungen glaubt der (meist private) Produzent der Daten, daß ihm auch diese Daten gehören - eine Art von Copyright.

Und sie werden noch juristisch in diesem Fehldenken bestärkt, weil in den entsprechenden Werkverträgen kein Hinweis darüber zu finden ist, daß die vom (öffentlichen) Auftraggeber mit Steuermitteln finanzierten Daten Eigentum des Auftraggebers sind und nicht des Auftragnehmers!

Ein letzter Punkt zu diesem Komplex: die **Datengläubigkeit**, zumindestens von offiziellen oder amtlichen Daten. Viele Menschen glauben fest daran, daß veröffentlichte (offizielle oder amtliche) Daten und Informationen per se richtig sind. Weit gefehlt! Das Beispiel der Abbildung 24 zeigt, daß auch (und nicht so selten) offizielle Statistiken falsch sein können, weil sie eben auch von Menschen gemacht wurden - und keinen Göttern, die allein unfehlbar sind.

Dieses „Mißverständnis“ führt mitunter soweit, daß - obwohl jeweils aktuelle wissenschaftliche Untersuchungen und Erhebungen zu neuen Resultaten kommen - diese Erkenntnisse zum Teil öffentlich nicht genutzt werden (dürfen), nur, um die „in Stein gehauenen“ Werte nicht anzutasten.

Diese Medaille hat aber auch zwei Seiten:

- Nicht nur, daß „alte“ Daten dadurch zu langsam ersetzt werden,
- nein, auch der Datenproduzent läuft immer mit der „Angst“ herum, seine in „Stein gehauene“ Produkte sind nicht valid genug, um dieser Prozedur gerecht zu werden.

In dieser Hinsicht sollten alle Daten-Beteiligten flexibler werden.

3.5.2. Informationsfrust - „Was ich nicht weiss, macht mich nicht heiss!“

Die im vorhergehenden Kapitel genannten Aspekte, aber auch

- die Angst vor der Datenflut und deren Bearbeitbarkeit,
- schlechte Erfahrungen, die mit Daten gemacht wurden, und
- daraus resultierendes Mißtrauen

verleiden nicht wenige Personen, von wichtigen Informationen und Daten immer mehr Abstand zu nehmen und sich als **daten-immun** zu deklarieren.

Das Folgedenken besteht in den Devisen:

- „Was ich nicht weiß, macht mich nicht heiß!“ bzw.
- „Der Mensch denkt, Gott lenkt.“

Die Folge solcher „Strategien“ ist ein **Informations-Analphabetismus**, der populistischen Informationskampagnen Vorschub leistet und die oben genannten negativen Erscheinungen geradezu verstärkt.

3.5.3. Informationsschutz - Datenschutz

Über die beiden Bereiche, die in den Kapiteln 3.5.1 und 3.5.2 behandelt wurden, wirkt als Dach vor „Daten-Unwetter“ der Datenschutz.

Das gilt insbesondere für

- alle personengebundenen Daten,
- alle wirtschaftlichen Daten, aus denen einzelbetriebliche Detailinformationen herausgelesen werden können,
- alle Sicherheitsdaten.



Die berechtigten Belange des Datenschutzes dürfen jedoch nicht dazu benutzt werden, nichtschutzbedürftige, aber wichtige Daten unter dem Mäntelchen des Datenschutzes unzugänglich zu machen.

Andererseits ergeben sich aber auch aus der Sicht der Personenkreise, die sich mit der Durchführung von Verkehrsverhaltensbefragungen befassen, in der Vorbereitungsphase und in der Feldphase nicht selten Probleme, die aus dem Datenschutz resultieren und deren Lösung sehr sorgfältig und gewissenhaft angegangen werden muß, da sonst die Ergebnisse der Untersuchung erheblich verzerrt werden.

In der Vorbereitungsphase spielt der Datenschutz eine besondere Rolle, gilt es doch hierbei die Grundlagen für seine Einhaltung und Gewährleistung zu schaffen und zu sichern. Die technischen und organisatorischen Maßnahmen konzentrieren sich dabei bezüglich der Einhaltung und Absicherung des Datenschutzes auf die folgenden Anforderungen: Zugangskontrolle, Abgangskontrolle, Speicherkontrolle, Benutzerkontrolle, Zugriffskontrolle, Übermittlungskontrolle, Eingabekontrolle, Auftragskontrolle, Transportkontrolle und Organisationskontrolle.

Die sachgemäße und zugleich im Sinne des Datenschutzes anforderungsgemäße Handhabung dieser Tätigkeiten erfordert Sorgfalt, Erfahrung und ein entsprechendes Verantwortungsbewußtsein. Sie bilden die technischen und organisatorischen Grundlagen zur Gewährleistung des Datenschutzes und sind in jedem Fall wie alle Datenschutzaufgaben genau einzuhalten.

Ein weiterer wichtiger Bereich, der bei der Vorbereitung von Verkehrsverhaltensbefragungen eine wichtige Rolle spielt, sind inhaltliche Maßnahmen.

Solche Punkte, deren Schwierigkeit und Komplexität mitunter anfangs unterschätzt wird, sind unter anderem:

- Datenschutz-Kompetenzen,
- Freiwilligkeit - Schriftlichkeit (bei schriftlichen Befragungen),
- Löschen der Adreßdateien und

- Trennung der Namens- und Adreßdaten von den Verkehrserhebungsdaten.

4. WEITERE ZUSAMMENHÄNGE ZWISCHEN DATENQUALITÄT UND INFORMATION

Im Kapiteln 2 und 3 wurden die Grundlagen und grundlegende Zusammenhänge zwischen den Bereichen der Datenqualität einerseits und der Information andererseits ausgeführt.

Diese Zusammenhänge schöpfen jedoch das Thema bei weitem nicht aus:

- In diesem Kapitel werden nun **weitere** Verbindungen zu dieser Fragestellung gebracht.
- Im Unterschied zu den genannten Kapiteln 2 und 3, in denen eher eine - für den Rahmen eines Vortragsmanuskripts - umfangreichere Darstellungsform gewählt wurde, werden die nun folgenden Gedanken **nur angetippt und skizzenhaft ausgewiesen**, da sonst der Rahmen vom mengenmäßigen Umfang gesprengt würde.

Ich bitte dafür um Verständnis.



4.1. Daten und Informationsgewinn

Zu Daten brauchen wir nicht nur Informationen, Daten liefern nicht nur selbst Information, wir müssen nicht nur über Daten Information haben, sondern Daten müssen auch weiterverarbeitbar sein und damit einen zusätzlichen Informationsgewinn ermöglichen.

Dieser Themenstellung ist der folgende Bereich „Daten und Informationsgewinn“ gewidmet.

4.1.1. Erkenntnis-(Modelle) und Information

Information hat im Sinne des Bereichs „Daten und Informationsgewinn“ etwas mit Erkenntnis zu tun.

Erkenntnis ist das Erfassen der Wahrheit einer Sache, die vom Erfassenden unabhängig besteht. In jedem Erkenntnisakt stehen sich der Erkennende (Subjekt) und das zu Erkennende/Erkannte (Objekt) gegenüber, wobei das Subjekt ein dem Objekt möglichst genau entsprechendes Abbild erzeugen soll.

Hilfsmittel zur Umsetzung dieses Prozesses sind sogenannte Erkenntnis-Modelle.

Die folgende Abbildung 25 gibt einen Überblick über solche „Abbildungsmaschinen“.

4.1.2. Verkehrsmodelle und Information

Spezifische Erkenntnis-Modelle sind nun die in der Verkehrsforschung und -planung eingesetzten Verkehrsmodelle ⁴³. Sie haben einen sehr vielfältigen Bezug zum Thema „Daten und Informationsgewinn“. Am wichtigsten scheint mir jedoch dabei die Kalibrierungsproblematik zu sein. Sie ist nicht der aber ein Schlüssel zur Validierung von Modellen (jeglicher Art). Dazu ist wichtig, daß

- wirklich kalibriert wird, denn dieser wesentliche Schritt wird häufig „vergessen“,
- auch dann eine Kalibrierung durchgeführt wird, wenn die Modellparameter von anderen bereits kalibrierten Modellen übernommen werden,
- Kalibrierung nicht mit Parameter-Eichung verwechselt wird:
 - Es können nämlich
 - * nichtvalide Modelle
 - * (richtig) geeicht werden,
 - was zu verdeckten aber erheblichen Fehlern führt.

4.1.3. Komplexität von Modellen

Modelle können zur Abbildung von Sachverhalten mehr oder weniger komplex sein - insbesondere die Verkehrsmodelle. Dazu ist folgendes zu sagen:

Problem	Erkenntnis system a la	Charakteristikum	Lösungsbeiträge durch
strukturiert	LOCKE	Übereinstimmung v. Experten aufgrund bekannter Daten	Datenbankabfrage u. Abstimmungsprozeß zwischen den beteiligten Modellbildnern (z.B. Delphi, s. CRAEMER 1977)
	LEIBNITZ	Problembeschreibung durch Modelle u. Axiome; alternat. Durchrechnung unter verschiedenen Annahmen	Modellrechnung, formale Herleitung od. Beweis
hinterhältig	KANT	alternative, sich ergänzende Problembeschreibungen; Probleme m. Datenerhebung u. Modellbildung	Einzelmodelle, erstellt durch Manager der betriebl. Hauptfunktionen Beschaffung, Fertigung, Vertrieb u. Verwaltung; Einsatz multivariater Verfahren, wie Faktorenanalyse und multidimensionaler Skalierung
	HEGEL	widersprüchliche Problembeschreibungen	Konflikt-Feld-Modell und dessen Erprobung im Planspiel (s. CRAEMER 1976)

Abbildung 25: Erkenntnis-Modelle ⁴³

⁴³ CRAEMER D.: Mathematisches Modellieren dynamischer Vorgänge. B.G.Teubner, Stuttgart 1985

⁴⁴ siehe zum Beispiel: DORFWIRTH J.-R., HERRY M.: Verkehrsmodell Österreich - Methode und mathematisches Grundmodell. Bundesministerium für Bauten und Technik, Straßenforschung, Heft 144, Wien 1980

- Mit Zunahme der Komplexität steigt nicht unbedingt die Genauigkeit der Ergebnisse,
- im Gegenteil, wenn die dafür erforderlichen Daten nicht vorhanden sind, kann sie sogar sinken.
- Daraus geht unmittelbar hervor, daß sich mit steigender Komplexität die Datenanforderungen erhöhen.
- Zunahme der Komplexität von Modellen bedeutet nicht unbedingt deren Erhöhung der Validität - das sind zwei verschiedene Paar Schuhe.
- Hohe Komplexität täuscht mitunter eine Wissenschaftlichkeit vor, die unter Umständen gar nicht vorhanden ist, was vor allem bei den rein mathematischen (Black-Box)-Modellen meistens der Fall ist.

Komplexität hat also etwas mit

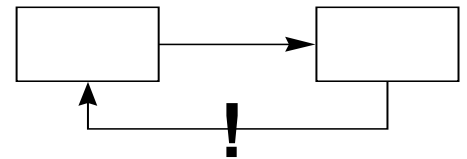
- Aufwand,
- Kenntnis,
- Erfahrungen,
- Sorgfalt und
- „Ehrlichkeit“

zu tun.

Neuere Forschungen verbinden Komplexität von Systemen auch mit der Chaos-Theorie (siehe Kapitel 4.4.2) ⁴⁵.

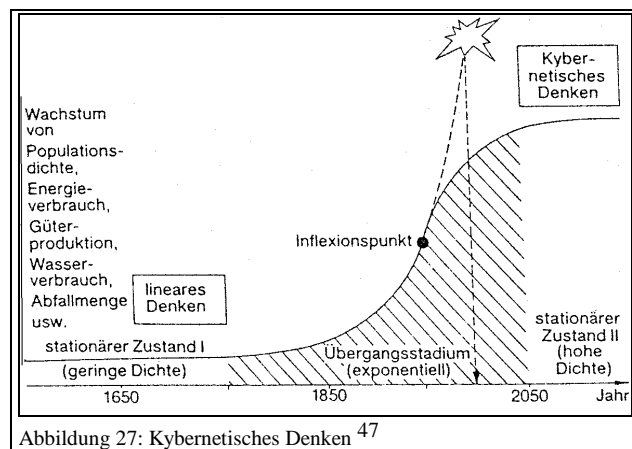
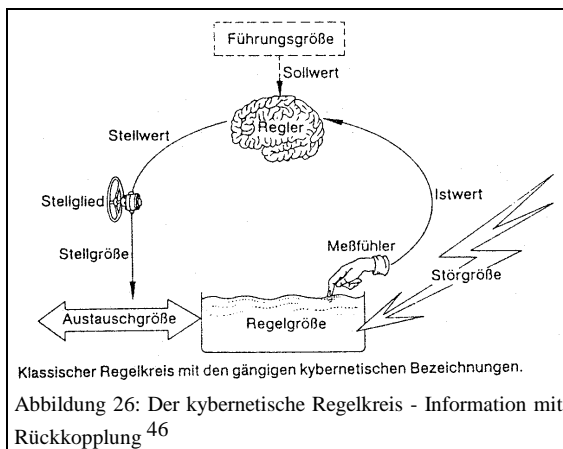
4.2. Informationsrückkopplung

Eines der entscheidendsten Entwicklungen in dem Zusammenhangsbereich „Datenqualität und Information“ ist die Informationsrückkopplung, weil erst durch sie wesentliche Entwicklungen möglich wurden und werden.



4.2.1. Kybernetischer Regelkreis

Grundlage für die Informationsrückkopplung ist der (kybernetische) Regelkreis: siehe Abbildung 26: Der kybernetische Regelkreis - Information mit Rückkopplung . Eng damit verflochten ist das kybernetische Denken - siehe Abbildung 27.



4.2.2. Modelle mit Informationsrückkopplung

Modelle mit Informationsrückkopplung sind sehr vielfältig, kommen in fast allen „Abbildungs“-Bereichen vor und reichen bis zu den Modellen der Chaos-Theorie (siehe Kapitel 4.4.2).

4.2.2.1. Delphi-Verfahren

Das Delphi-Verfahren besteht in einem mehrstufigen Bewertungsverfahren, das in der Verkehrsplanung oft angewendet wird:

- Die erste Phase besteht in einer (Anfangs-)Bewertung,
- in der zweiten Phase werden die Bewertungsergebnisse der ersten Phase bekanntgegeben, ausgewertet und diskutiert.
- In der dritten Phase wird erneut bewertet.
- Dieser Prozeß kann sich wiederholen.

4.2.2.2. Retransferenz

Bei der Einführung von Road Pricing tritt ein sogenannter **Rücklauf-Effekt** auf ⁴⁸:

⁴⁵ PEAK D., FRAME M.: Komplexität - das gezähmte Chaos. Birkhäuser-Verlag Basel - Boston - Berlin, 1995

⁴⁶ FLECHTNER H.J.: Grundbegriffe der Kybernetik. Deutscher Taschenbuch Verlag, 1984

⁴⁷ VESTER F.: Neuland des Denkens. Deutscher Taschenbuch Verlag, 1986

Mit steigender Maut wächst - verständlicherweise - zum Beispiel die Anzahl der Routenänderungen, d.h. das Ausweichen auf das parallel zur Mautstrecke führende Straßennetz. Ist nun diese Routenverlagerung sehr hoch, so entschließt sich - erwartungsgemäß - ein Teil der Routenverlagerer, aufgrund der Verkehrsqualitätsverschlechterung von der parallelführenden mautfreien Strecke wieder zurück auf die bemaute Autobahn zu wechseln.

Diesen Prozeß nennen wir Rückverlagerung: Analog trifft er für andere Verkehrsverhaltensänderungen zu.

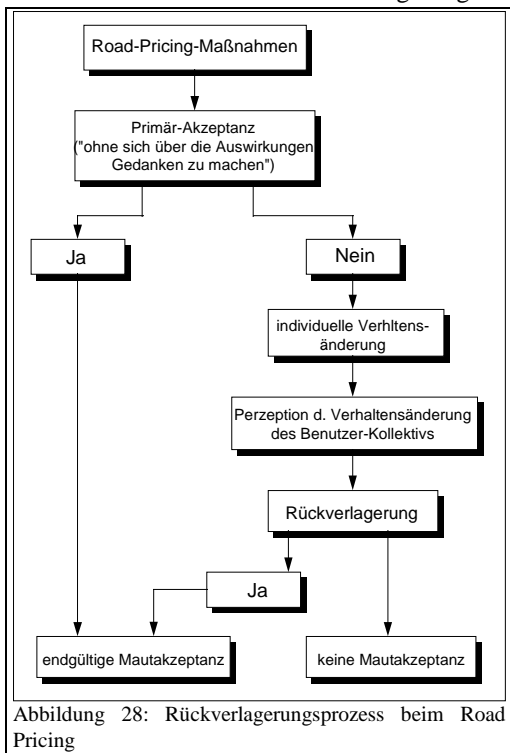


Abbildung 28: Rückverlagerungsprozess beim Road Pricing

Das Grundsche ma des Rückverlagerungsprozesses bei Mautakzeptanzentscheidungen ist in der folgenden beiden Abbildungen dargestellt. Mautuntersuchungen haben diese Rückverlagerungsprozesse nicht nur zu berücksichtigen, sondern auch zahlenmäßig zu erfassen.

Dies ist nach Auffassung des Autors bisher nur in der Arbeit HERRY / SNIZEK 1992 geschehen ⁴⁹. In der Abbildung 30 ist ein konkretes Beispiel für Retransferenz ausgewiesen.

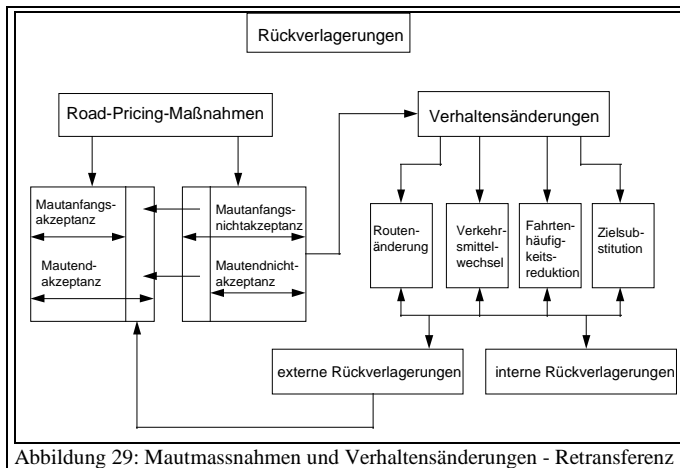


Abbildung 29: Mautmassnahmen und Verhaltensänderungen - Retransferenz

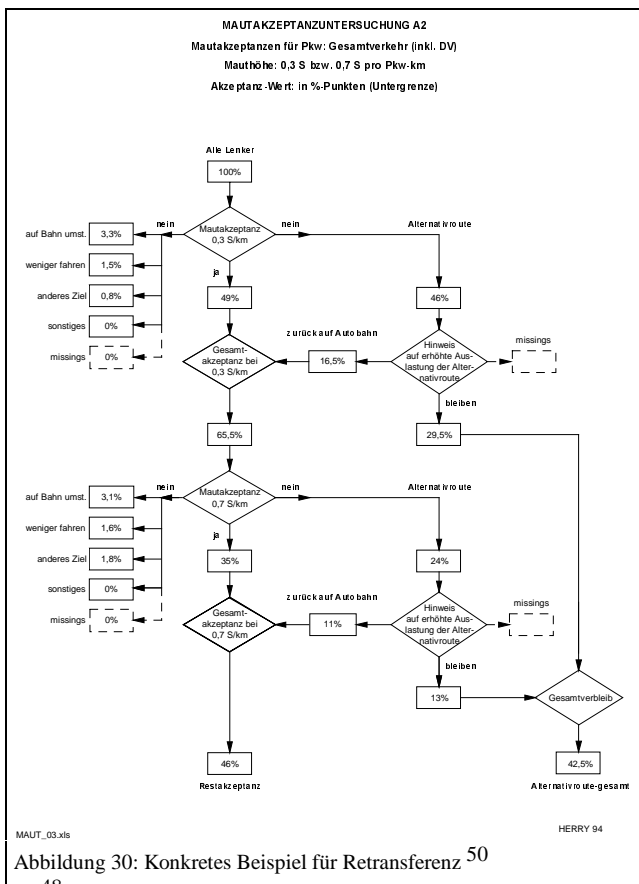
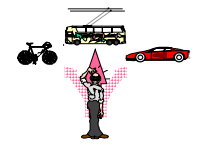


Abbildung 30: Konkretes Beispiel für Retransferenz ⁵⁰



4.2.2.3. Manipulierte Informationsrückkopplung

Es gibt auch manipulierte Informationsrückkopplung: So wurde im österreichischen Fernsehen anlässlich der Einführung der Nachtbuslinien in Wien eine Zuseher-Befragung durchgeführt. Nachdem erste Resultate vorhanden waren, wurden diese bekanntgegeben, gleichzeitig wurde jedoch dazu aufgerufen, daß jener Personenkreis, dessen Meinung zum Zeitpunkt dieser Bekanntgabe noch in der Minderheit war, stärker in Erscheinung treten solle, um eventuell das Resultat noch „umdrehen“ zu können! Eine solche Informationsrückkopplung ist natürlich nicht wünschenswert.

⁴⁸ HERRY M.: Toll Acceptance and Retransference. Proceedings of the 23rd PTRC European Transport Forum, Warwick 1995
⁴⁹ HERRY M., SNIZEK S.: Road Pricing in Österreich mit Informationsrückkopplung (Road-Pricing-II-Studie). Im Auftrag des Bundesministeriums für wirtschaftliche Angelegenheiten, Wien 1992
⁵⁰ HERRY M.: Mautakzeptanzverhalten auf der A2. Wien 1995

4.3. Information und Entscheidungsverhalten

Information hat im Sinne des Bereichs „Datenqualität und Information“ hat auch etwas mit Entscheidungen zu tun.

4.3.1. Entscheidungsverhalten

Wichtige Begriffe zum Entscheidungsverhalten sind ⁵¹:

- Wahl zwischen Alternativen:
 - Rationalität:
 - * Entscheidungsergebnis
 - * Entscheidungsträger
 - Bewußtheit:

* Entscheidungsprozeß

- Selbstverpflichtung:
 - Bindung an die Entscheidung

Das **Phasenschema der Entscheidung** ist in der Abbildung 31 dargestellt. Die **Elemente eines Entscheidungssystems** zeigt Abbildung 32.

4.3.2. Information und Kooperation

Eng mit dem Entscheidungsverhalten ist der Begriff der Kooperation verbunden und in Ihrem Umfeld die Information.

Das Grundproblem der Kooperation tritt dann auf, wenn die Verfolgung des Eigeninteresses durch jeden einzelnen „Partner“ zu einem schlechten Ergebnis für alle „Beteiligten“ zusammen führt. Um bei der Untersuchung der enormen Menge spezifischer Situationen, die diese Eigenschaft besitzen, voranzukommen, ohne sich zu sehr in den Details einzelner Situationen zu verlieren, ist eine geeignete Darstellung der gemeinsamen Merkmale dieser Situationen erforderlich ⁵². Glücklicherweise existiert diese in Form des berühmten **Gefangenendilemma-Spiels**:

Das Gefangenendilemma ist ein Spiel mit zwei Spielern, von denen jeder zwei Entscheidungsmöglichkeiten hat, nämlich zu kooperieren oder nicht zu kooperieren. Nichtkooperation wird Defektion genannt. Jeder muß seine Wahl treffen, ohne zu wissen, wie der andere sich verhalten wird. Unabhängig vom Verhalten des jeweils anderen führt Defektion zu einer höheren Auszahlung als Kooperation. Das Dilemma liegt darin, daß es für jeden Spieler, unabhängig vom Verhalten des anderen, vorteilhafter ist, zu defektieren, daß jedoch beiderseitige Defektion für jeden Spieler ungünstiger ist als wechselseitige Kooperation!

Das Gefangenendilemma wurde ungefähr 1950 von Merrill Flood und Melvin Dresher erfunden und kurz darauf von A. W. Tucker formalisiert. Situationen, die über paarweise Interaktionen hinausgehen, können mit dem komplexeren N-Personen-Gefangenendilemma modelliert werden ⁵³.

AXELROD macht nun Vorschläge für ein erfolgreiches Verhalten: Es handelt sich um vier einfache Vorschläge, wie man in einem dauerhaften iterierten Gefangenendilemma gut abschneidet ⁵⁴:

1. Sei nicht neidisch.
2. Defektieren nicht als erster.
3. Erwidere sowohl Kooperation als auch Defektion.
4. Sei nicht zu raffiniert.

Das könnten auch gute Ratschläge für die Kooperation von Verkehrs- und Raumplanern sein!

4.4. Information und Bestimmtheit

Zum Thema „Datenqualität und Information“ gehört natürlich auch der Bereich der Bestimmtheit im allgemeinen und ihren Anwendungen im besonderen. Bezüglich der Bestimmtheit unterscheiden wir

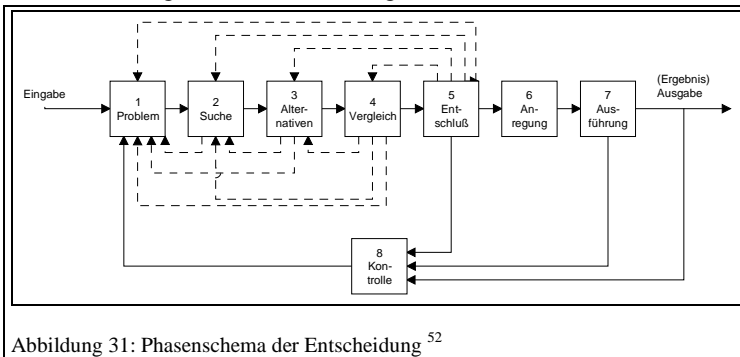


Abbildung 31: Phasenschema der Entscheidung ⁵²

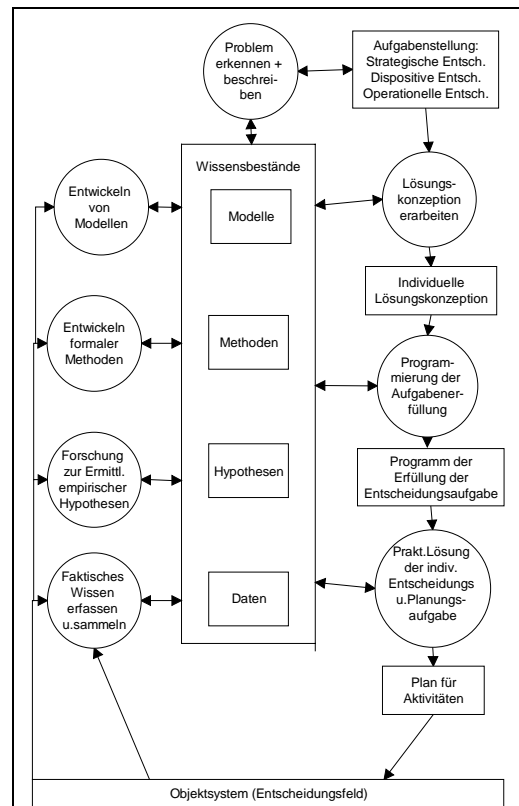


Abbildung 32: Elemente eines Entscheidungssystems ⁵²

⁵¹ SZYPERSKI N., WINAND U.: Entscheidungstheorie. Poeschel-Verlag, 1974

⁵² AXELROD R.: Die Evolution der Kooperation. R. Oldenbourg Verlag München, 1988

⁵³ OLSON, M. Jr.: Die Logik des kollektiven Handelns. Tübingen: Mohr

- deterministische und
- stochastische Systeme.

4.4.1. Simulationen

Unter (mathematischer) Simulation wird allgemein ein Verfahren zur Durchführung von Experimenten unter Zuhilfenahme eines Computers und Benutzung mathematischer Modelle mit dem Ziel verstanden, Aussagen über das Verhalten eines (realen) Systems zu gewinnen.

Es handelt sich dabei ⁵⁴

- um keine feststehende Methode, sondern die Vorgehensweise richtet sich nach der Problemstellung und kann sehr unterschiedlich ausfallen,
- um das Studium des Systemverhaltens. Es werden also Systeme analysiert und mit mathematischen Methoden nachvollzogen.

4.4.2. Information und Chaos



Ein zentraler Aspekt unserer Betrachtungen waren Systeme mit Informationsrückkopplung (siehe Kapitel 4.2), andererseits unterschieden wir am Anfang des Kapitels 4.4 deterministische und stochastische Systeme, wobei sie - in der „nichtchaotischen“ Denkweise stets relativ voneinander unabhängig betrachtet wurden: deterministische Systeme unterliegen deterministischen funktionalen Zusammenhängen und stochastische Systeme stochastischen funktionalen Zusammenhängen.

Ein **Brückenschlag** zwischen beiden Systemen gelang den Begründern der Chaos-Theorie **über die Informationsrückkopplung:**

Nach den Studien von RUELLE und TAKENS ⁵⁵ und den Experimenten von GOLLUP und SWINNEY ⁵⁶ in den 70-iger Jahren ergab sich ein Wendepunkt in der Betrachtungsweise über das Einsetzen von Turbulenzen.

Ein neues Paradigma entstand, und es erhielt den Namen „CHAOS“ von Jim YORKE, einem angewandten Mathematiker ⁵⁷.

Was wir heute Chaos nennen, ist eine „Zeitentwicklung“ mit empfindlicher Abhängigkeit von den Anfangsbedingungen.

Die Bewegung auf seltsamen Attraktoren ist demnach chaotisch ⁵⁸. Man spricht auch von deterministischem Rauschen, wenn die irregulären Oszillationen als **zufällig** erscheinen, aber der Mechanismus, der sie produziert, **deterministisch** ist.

Die Geschichte der Wissenschaft ist voll von Ironie. So ist das Werk von FATOU und JULIA, das den jungen MANDELBROT wegen des fehlenden geometrischen Inhalts von der reinen Mathematik abgebracht hat, als eine zentrale Anwendung der Fraktale auf Hauptgebiete der Mathematik wieder aufgetaucht und wird wegen der bildlichen Schönheit der Fraktale von weiten Kreisen begeistert aufgenommen.

Gaston JULIA war ein Student von POINCARÉ und beschäftigte sich mit Iterationen von Abbildungen der komplexen Zahlenebene, zum Beispiel $z \rightarrow z^2 + c$, wobei c eine Konstante ist. Dahinter steckt die Idee, einen Wert für c zu fixieren und zu fragen, was mit einem gegebenen Anfangswert z passiert, wenn diese Formel iteriert wird.

Oberflächlich gesehen läßt sich ein Hauptunterschied feststellen. Einige Anfangswerte z bewegen sich rasch ins Unendliche, die restlichen nicht.

Wenn sich ein Punkt bei der Iteration ins Unendliche bewegt, malt man ihn schwarz an, andernfalls weiß. Hierdurch stellt man das „Einzugsgebiet“ des Punktes im Unendlichen dar.

Die JULIA-Menge ist sein Rand.

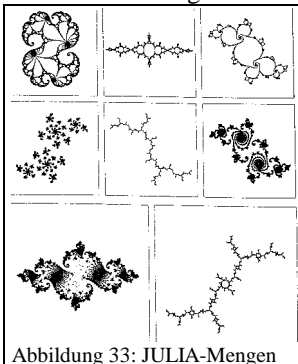


Abbildung 33: JULIA-Mengen

In chaotischen Systemen erzeugt also die deterministische Ordnung die Unordnung des Zufalls - und das mittels **einfacher rückgekoppelter Schleifen!**

Der Prototyp einer solchen Rückkopplung (mit bestimmten Anfangswerten) ist der **Lebkuchenmann** - auch MANDELBROT-Menge genannt:

Dazu iterieren wir zu einer gegebenen komplexen Zahl c_0 die Abbildung $z \rightarrow z^2 + c_0$ für (alle) komplexen Zahlen z und erhalten die JULIA-Menge für c_0 (siehe oben). Ist diese Menge (topologisch) *zusammenhängend*, so ist c_0 ein Element (Punkt) der MANDELBROT-Menge, ist sie nicht (topologisch) *zusammenhängend*, so liegt c_0 außerhalb der MANDELBROT-Menge.

⁵⁴ HERRY M., ZUMKELLER D.: Simulation von Verkehrssystemen. In: Fortschritte in der Simulationstechnik, 6. Symposium zur Simulationstechnik in Wien 1990, Tagungsband, Wien 1990

⁵⁵ RUELLE D., TAKENS F.: On the Nature of Turbulence. Commun.Math. Phys., 20, 167-192, 1971

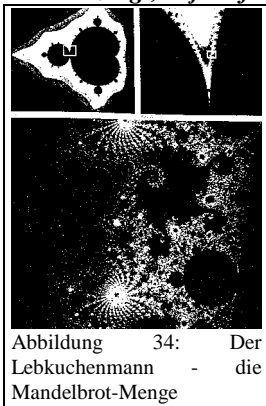
⁵⁶ GOLLUP J.P., SWINNEY H.L.: Onset of Turbulence in a rotating Fluid. Phys. Rev. Lett., 35, 927-930, 1975

⁵⁷ LI T., YORKE J.A.: Period three implies Chaos. Amer. Math. Monthly, 82, 985-992, 1975

⁵⁸ RUELLE D.: Chance and Chaos. Princeton University Press, 1991

Entscheidend für unser Thema „Datenqualität und Information“ scheint mir nun die Existenz von sogenannten (seltsamen) Attraktoren⁵⁹, die wie Faktoren (als Ergebnis einer Faktorenanalyse) im „Hintergrund tieferer“ Zusammenhänge erklären.

Ich bin nun überzeugt, daß es für die chaostheoretischen Bestandteile der wichtigsten Mobilitäts-„Bilder“ solche seltsamen Attraktoren gibt.



Ich bin überzeugt, daß dieser Zugang die Erklärung von Mobilität, die im gegenwärtigen Stadium auf der Stelle tritt, wesentlich befruchten würde und eine entscheidende Weiterentwicklung dieses wichtigen Gebietes mit Hilfe der Erkenntnisse der Chaostheorie darstellen könnte.

4.5. Formen von Informationen - Datenformen

Abschließend wollen wir das Thema „Datenqualität und Information“ durch einige kurze Anmerkungen zu den Formen von Informationen beenden.

4.5.1. Individual-Information und Kollektiv-Information



In der Regel wird bei Berechnungen von Kollektivwerten zu verkehrlichen Größen (zum Beispiel Kraftstoffverbrauch) von Einzelwerten (zum Beispiel Durchschnittsverbräuchen von „repräsentativen“ Einzelfahrzeugen) ausgegangen.

Dabei wird jedoch außer acht gelassen, daß^{60 61}

- die Herleitung der Kollektivwerte nicht einfach aus der „Summe“ der Einzelwerte gebildet werden kann, da die entsprechenden funktionalen Zusammenhänge in der Regel
 - weder linear
 - noch „symmetrisch“ verlaufen,
- die Kollektivwerte gegenüber den Einzelwerten in der Regel
 - nicht mittelwertsinvariant und
 - schon gar nicht streuungsinvariant sind.

Mit aggregierten Berechnungsmethoden sind diese Probleme nur sehr schwer zu lösen.

Besser wäre in diesem Fall eine Simulationsrechnung (siehe Kapitel 4.4.1).

4.5.2. Aggregation und Disaggregation von Daten

Der Aggregations- bzw. Disaggregationsgrad von Daten ist ein entscheidendes Merkmal für ihre Beziehungen zu den Bereichen „Information“ und „Datenqualität“:

- Die meisten Daten - mit auch der höchsten Datenqualität - sind weder sehr aggregiert noch sehr disaggregiert, sie kommen aus einer mittleren Aggregationsebene und entstammen unseren unmittelbaren Lebens- und Erlebnisbereichen.
- Globale Daten wie auch sehr differenzierte Aussagen verlangen umfangreiche Datenerhebungen und Berechnungen: So ist zum Beispiel die „einfache“ Frage nach der Anzahl aller Wege der Personen in Österreich im Laufe eines Jahres nur sehr schwer zu beantworten.

Aber gerade solche Informationen haben einen sehr großen Informationsgehalt und sind von strategischer Bedeutung: Zum Beispiel ist es für Maut-Vignetten-Berechnungen wichtig, zu wissen, wieviel Mal ein und dasselbe ausländische Fahrzeug pro Jahr nach Österreich einfährt.

Überhaupt sind Fahrzeug- oder Personen-gebundene Verkehrsaussagen zwar sehr wesentlich, aber auch sehr schwierig zu ermitteln.

Sie sind wichtiger als sehr disaggregierte Informationen.

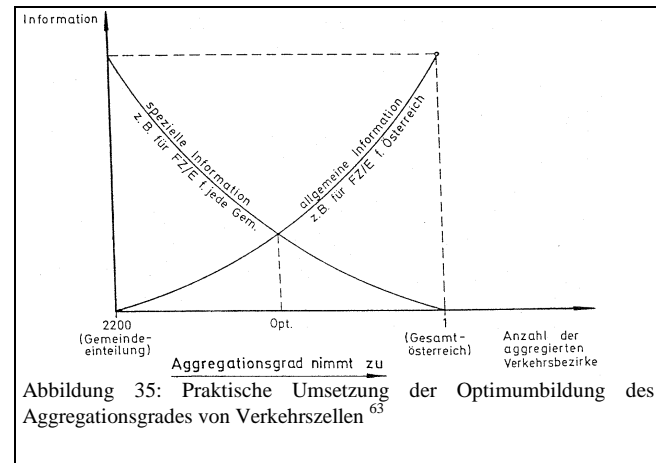
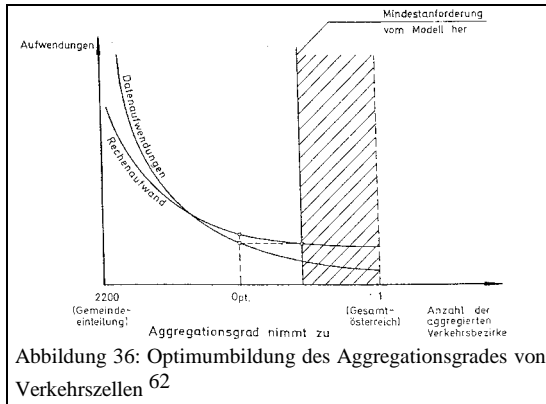
Auf ihre Ermittlung sollte in Zukunft mehr Augenmerk gelegt werden.

⁵⁹ LEWIN R.: Die Komplexitätstheorie - Wissenschaft nach der Chaosforschung. Verlag Hoffmann und Campe, Hamburg, 1993

⁶⁰ HERRY M.: Methode und mathematisches Grundmodell zu Berechnungen im Mengengerüst von Entscheidungshilfen. Diss. An der TU Wien, Wien 1982

⁶¹ DORFWIRTH J.R., HERRY M.: Nutzen-Kosten-Untersuchung im Straßenbau - Berechnung des Mengengerüsts von Straßennetzen. Bundesministerium für Bauten und Technik, Straßenforschung, Heft 151, Wien 1979

- Bezüglich der Aggregationen von Verkehrszellen und den entsprechenden Informationen bzw. Rechenaufwendungen gelten die Darstellungen der folgenden beiden Abbildungen.



5. SCHLUSSFOLGERUNGEN

Folgende Schlußfolgerungen können aus den Darlegungen gezogen werden:

- Das Thema ist vor allem deshalb wichtig, weil
 - einerseits, die Information zum wesentlichen Bestandteil unseres heutigen Lebens geworden ist (Informationsgesellschaft) und auch eine der Grundlagen für die „Bestimmung“ und Gestaltung **unserer** Zukunft sein wird,
 - andererseits, aber gerade dazu - und das vor allem in der Verkehrsplanung - nicht nur Daten an sich, sondern Daten von hoher Qualität gebraucht werden.
- **Wichtig** ist nun,
 - nicht nur zu wissen, auf **welchen** Informationen die davon erzeugten Daten aufbauen, sondern auch - und vor allem -
 - welche dieser Informationen **unvollständig** sind,
 - und welche Informationen dazu **überhaupt nicht** zur Verfügung stehen.
Der zweite Bereich bedingt eine Non-Reported-Analyse der dritte eine Non-Response-Analyse, Verfahren, die leider allzu oft bei der Datengewinnung außer acht gelassen oder zumindest unterschätzt werden.
- Der Prozeß der **Rückkopplung** spielt im Bereich der Informationstheorie und -praxis eine enorm wichtige Rolle. Durch sie konnten erst solche bedeutende Bereiche, wie
 - (kybernetische) Regelkreise oder
 - auch die Chaos-Theorie entwickelt werden.
- Die Beziehung „Daten - Information“ erzeugt also vier wesentliche Zusammenhangsbereiche:
 - Information --> Daten (*Eingangsinformationen* für Datenerstellung)
 - Daten als Information (*Dateninhalt*)
 - Daten --> Information (*Informationsgewinnung* aus Daten(-Weiterverarbeitung) - *Ausgangsinformation*)
 - Information über die Daten (*Informationszugang*)
- Datenqualität hat etwas mit **Messen** und folglich mit dem **Meßfehler-Problem** zu tun.
- Meßfehler von empirischen Daten setzen sich aus stichprobenbedingten und Design-Fehlern zusammen, wobei letztere zu wenig Beachtung finden, was in der Regel zu invaliden Daten führt!
- Die Meßfehler der analytischen Daten werden so gut wie gar nicht verfolgt: Die Fehlerfortpflanzung ist ein Fremdwort in der praktischen Verkehrsplanung.
- Es ist sehr wichtig, den Genauigkeitsgrad für die jeweilige Datengewinnung zu bestimmen, um gezielt, sparsam und effizient Daten zu produzieren. Dabei ist wichtig, wofür die Daten gebraucht werden.
- Das Datenverarbeitung mit unvollständigen oder fehlenden Informationen wird in der Verkehrsplanung größtenteils auf relativ niedrigem Niveau - wenn überhaupt - abgewickelt.
Auf diesem Gebiet sollten vermehrt Anstrengungen unternommen werden, um diese gefährliche Lücke zu schließen.

⁶² DORFWIRTH J.-R., HERRY M.: Verkehrsmodell Österreich - Methode und mathematisches Grundmodell. Bundesministerium für Bauten und Technik, Straßenforschung, Heft 144, Wien 1980

- Daten und Informationen spiegeln nicht nur die Wirklichkeit, sondern auch Einstellungen, Meinungen und Akzeptanzen wider. Diese müssen verstärkt erforscht werden, um stringenter Verkehrsverhaltensklärungen erforschen zu können.
- **Wichtig** erscheint mir, relativ „zentral“ zu klären,
 - **welche Information/Daten** wirklich und zu welchem Zweck zukünftig **gebraucht** werden,
 - **wie genau** und umfangreich diese Information/Daten sein müssen und
 - **wer sie produzieren und erhalten** soll.
- Mit Statistik wird viel gelogen (bewußt oder unbewußt). Wir sollten uns bemühen, dieses Ausmaß empfindlich zu reduzieren.
- Zu viele Daten zu haben, ist genau so schlecht, wie zu wenig. Die Zukunft verlangt dazu mehr Ausgewogenheit.
- Daten und Informationen müssen stärker als bisher **harmonisiert** werden.
- Auftraggeber von Datenerstellungen müssen auch die Besitzer dieser Informationen sein und die Daten - unter Berücksichtigung des Datenschutzes - allen Benutzern zur Verfügung stellen.
- In Österreich herrscht noch eine zu starke Datengläubigkeit von amtlichen Informationen. Wir sollten mithelfen, diese Situation zu normalisieren.
- Wir sollten stärker die bestehenden Verkehrsmodelle in Richtung von **Erkenntnismodellen** entwickeln.
- Die Informationsrückkopplung ist das stärkste Instrument, das wir in der Informationsforschung und -praxis zur Zeit haben. Es wird die Zukunft der Informationsgewinnung und -weiterverarbeitung sowie -veredelung eindeutig bestimmen.
- Die Bereiche Informationsbereitstellung und Entscheidungsverhalten klaffen noch weit auseinander und müssen in Zukunft stärker miteinander verknüpft werden.
- Entscheidend für unser Thema „Datenqualität und Information“ scheint mir die Existenz von sogenannten (seltsamen) Attraktoren zu sein, die wie Faktoren (als Ergebnis einer Faktorenanalyse) im „Hintergrund tiefere“ Zusammenhänge erklären.
- Ich bin überzeugt, daß es für die chaostheoretischen Bestandteile der wichtigsten Mobilitäts-„Bilder“ solche seltsamen Attraktoren gibt.
- Außerdem bin ich davon überzeugt, daß dieser Zugang die Erklärung von Mobilität, die im gegenwärtigen Stadium auf der Stelle tritt, wesentlich befruchten würde und eine entscheidende Weiterentwicklung dieses wichtigen Gebietes mit Hilfe der Erkenntnisse der Chaostheorie darstellen könnte.

6. LITERATUR

- ARMOOGUM J., MADRE J.-L.: Nonresponse Correction in the French 1993-94 NPTS: The Example of Daily Trips. INRETS, Paris, 1996
- AXELROD R.: Die Evolution der Kooperation. R.Oldenbourg Verlag München, 1988
- AXHAUSEN K.: Erhebungsverfahren im Verkehrswesen. Innsbruck 1996
- AXHAUSEN K., BOGNER W., HERRY M., VERRON H., VOLKMAR H., WICHMANN W., ZUMKELLER D.: Merkblatt zur Messung von Präferenzstrukturen - Methoden der „Stated Preferences“. 1996
- AXHAUSEN K., KÖLL H., HERRY M.: Long Distance Travel Demand Measurement Methods Pilot Study. Im Auftrag des Bundesministeriums für Wissenschaft, Verkehr und Kunst, Innsbruck, Wien 1996
- BAUER F.L., GOOS G.: Informatik. Springer-Verlag Berlin-heidelberg-New York, 1982
- BRÖG W., ERL E., MEYBURG A.H.: Problems of Non-reported Trips in Surveys of Non-home Activity Patterns. 61st Annual Meeting of the Transportation Research Board, Washington, D.C., January, 18-22, 1982
- CRAEMER D.: Mathematisches Modellieren dynamischer Vorgänge. B.G.Teubner, Stuttgart 1985
- DORFWIRTH J.-R., HERRY M.: Verkehrsmodell Österreich - Methode und mathematisches Grundmodell. Bundesministerium für Bauten und Technik, Straßenforschung, Heft 144, Wien 1980
- DORFWIRTH J.R., HERRY M.: Nutzen-Kosten-Untersuchung im Straßenbau - Berechnung des Mengengerüsts von Straßennetzen. Bundesministerium für Bauten und Technik, Straßenforschung, Heft 151, Wien 1979
- FLECHTNER H.J.: Grundbegriffe der Kybernetik. Deutscher Taschenbuch Verlag, 1984
- FORNEY, G. D., Jr.: Information Theory, course notes for EE-376. Stanford University, 1972
- HAUTZINGER H. et al.: Mobilität - Ursachen, Meinungen, Gestaltbarkeit. Studie im Auftrag des VDA, der BAG und des ADAC, Heilbronn 1994
- HERRY M.: Mobilität von Personen und Gütern. Vorlesung an der TU Wien, Wien 1996
- HERRY M.: EDV-Software - ein hilfreiches Instrument für die Verkehrsplanung? Vortrag zum Symposium CORP'96, Tagungsband, S.205-218, Kapitel 2, Wien 1996
- HERRY M.: Toll Acceptance and Retransference. Proceedings of the 23rd PTRC European Transport Forum, Warwick 1995
- HERRY M.: Österreichisches Verkehrs-Plafonierungsmodell. Im Auftrag des Bundesministeriums für öffentliche Wirtschaft und Verkehr, Wien 1995
- HERRY M.: Mautakzeptanzverhalten auf der A2. Wien 1995
- HERRY M.: Die methodischen Grenzen der Direkten Nutzenmessung. Kolloquium "Veränderungen im Verkehrsverhalten: Messung und Modellierung mit den Methoden der Direkten Nutzenmessung" am 10. März 1994 in Karlsruhe
- HERRY M.: Der Güterverkehr in Österreich. ÖROK-Schriftenreihe, Heft 100, Wien 1994
- HERRY M.: Wegekosten und Maut in Österreich. In: Schriftenreihe der Institute für Eisenbahnwesen sowie Straßenbau und Verkehrswesen der TU Graz, Heft 15, Graz 1993
- HERRY M.: Auswirkungen und Behandlung von Datenschutzproblemen bei haushaltsbefragungen zum Verkehrsverhalten. In: Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft, B87, 1987
- HERRY M.: Gewichtung der KONTIV 1982. In: Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft, B85, 1986
- HERRY M.: Methode und mathematisches Grundmodell zu Berechnungen im Mengengerüst von Entscheidungshilfen. Diss. An der TU Wien, Wien 1982
- HERRY M., ROSINAK W.: Vorher-Nachher-Untersuchung zur Einführung der Parkraumbewirtschaftung in den Bezirken 6 bis 9. Im Auftrag des Magistrats der Stadt Wien, Wien 1997

- HERRY M., SAMMER G.: Gewichtung und Hochrechnung der Österreichischen Verkehrsverhaltenshebung. Im Auftrag des Bundesministeriums für Wissenschaft, Verkehr und Kunst, Wien 1997
- HERRY M., SNIZEK S.: Road Pricing in Österreich mit Informationsrückkopplung (Road-Pricing-II-Studie). Im Auftrag des Bundesministeriums für wirtschaftliche Angelegenheiten, Wien 1992
- HERRY M., ZUMKELLER D.: Simulation von Verkehrssystemen. In: Fortschritte in der Simulationstechnik, 6. Symposium zur Simulationstechnik in Wien 1990, Tagungsband, Wien 1990
- INFRAS: Wie gau ist genau genug? Soziale Kosten und Nutzen im Verkehrsbereich. Essay-Papier im Auftrag des Dienstes für Gesamtverkehrsfragen der Schweiz, Zürich, 1993
- JOHANNESSON, R.: Informationstheorie. Verlag Addison-Wesley, Lund 1988
- KÖLTRINGER R.: Gültigkeit von Umfragedaten. Böhlau Verlag Wien - Köln - Weimar, 1993
- KRÄMER W.: So lügt man mit Statistik. Campus-Verlag Frankfurt - New York, 1995
- KRUG W., NOURNEY M.: Wirtschafts- und Sozialstatistik - Gewinnung von Daten. R.Oldenbourg Verlag München Wien, 1982
- LEWIN R.: Die Komplexitätstheorie - Wissenschaft nach der Chaosforschung. Verlag
- MASSEY J.L.: Information Theory: the Copernican System of Communications. IEEE Commun Mag. 22:26-28, 1984
- MIZEROVSKY H.: Kommunikation im Ernstfall. Wien - München - Zürich
- OLSON, M. Jr.: Die Logik des kollektiven Handelns. Tübingen: Mohr
- PEAK D., FRAME M.: Komplexität - das gezähmte Chaos. Birkhäuser-Verlag Basel - Boston - Berlin, 1995
- POLAK J.W., AMPT E.S.: An Analysis of Response Wave and Nonresponse Effects in Travel Diary Surveys. Paper presented at the 4th International Conference on Survey Methods in Transport, Steeple Aston, 9-11 September 1996
- RICHTER K.-J.: Der Entropieansatz in der Verkehrsanalyse und in der Verkehrsplanung. In: Österreichische Zeitschrift für Verkehrswissenschaft, Heft 1/90, S. 23-29
- SACHS L.: Angewandte Statistik. Springer-Verlag, 1992
- SAMMER G., FALLAST K., KÖSTENBERGER H. et al.: Verkehrsuntersuchung zum Nahverkehrskonzept Salzburg 1982. Im Auftrag der Salzburger Landesregierung et al., Graz 1985
- SHANNON C.E.: A Mathematical Theory of Communication. Bell Sys Tech J. 27:379-423 (Part I), 1948
- STEYER R., EID M.: Messen und Testen. Springer-Verlag Berlin Heidelberg New York, 1993
- SZYPERSKI N., WINAND U.: Entscheidungstheorie. Poeschel-Verlag, 1974
- VESTER F.: Neuland des Denkens. Deutscher Taschenbuch Verlag, 1986
- WATZLAWICK P.: Wie wirklich ist die Wirklichkeit. Piper-Verlag München - Zürich, 1978
- WIENER N.: Cybernetics. Paris, New York 1948