

Formulierung und Klassifikation von Fragen

Gallhofer, Irmtraud N.; Saris, Willem E.

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Gallhofer, I. N., & Saris, W. E. (2000). Formulierung und Klassifikation von Fragen. *ZUMA Nachrichten*, 24(46), 43-72.
<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-211156>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

FORMULIERUNG UND KLASSIFIKATION VON FRAGEN¹

IRMTRAUD N. GALLHOFER UND WILLEM E. SARIS

Häufig untersucht man, welchen Einfluß die Formulierung von Fragen auf Antworten hat, während das Problem der Übersetzung von Begriffen in Fragen kaum erörtert wird. In dieser Studie wollen wir uns mit der Übertragung von Konstrukten in Fragen befassen. Durch die Operationalisierung von Konstrukten kann angegeben werden, auf welche intuitiven Begriffe sie sich beziehen. Danach kann anhand der intuitiven Begriffe festgestellt werden, welche Fragen in Betracht kommen. Intuitive Begriffe beziehen sich auf kognitive Überzeugungen, Evaluationen, evaluative Überzeugungen, affektive Bewertungen, Präferenzen, Werte, Normen, Verhalten, Handlungstendenzen und dergleichen. Um den Zusammenhang zwischen den intuitiven Konzepten und ihrer sprachlichen Formulierung in den Items zu verdeutlichen, werden die Strukturen von Behauptungen analysiert. Wenn man die Begriffe mit der Frageform verknüpft, erhält man ein Verfahren zur systematischen Klassifizierung von Fragen, das sich auch eignet, um Fehler in Fragen zu entdecken, wodurch man eine erhebliche Verbesserung der Qualität der Fragen erzielen kann.

A great deal has been written about the effects of the wording of survey questions on the responses. On the other hand, relatively little attention has been given to the content of the questions or the translation of theoretical concepts into corresponding questions for survey research. In this paper we concentrate on the links between a set of basic concepts for social science research and questions which can be formulated in order to measure these concepts. The basic concepts we will discuss are: cognitions, evaluations, evaluative beliefs, feelings, preferences, rights, norms, policies, behavior, action tendencies, expectations and a number of demographic characteristics. In order to clarify the link between the concepts and their verbal expression (assertions), we analyze the structures of sentences expressing the different

¹ Die Niederländische Organisation für Wissenschaftliche Forschungen (NWO) ermöglichte diese Studie (Subventionsnummer 510-60-026). Weitere Informationen sind erhältlich bei der Stiftung für Soziometrische Untersuchungen, Meander 407, 1181 WN Amstelveen, Niederlande, Tel. x-31-20-64 36 418, E-Mail: Gallhof@lbn.net.

concepts. We also look at the specific types of questions which can be asked about such concepts and discusses their meaning. The combination of the classifications of concepts and questions provides a procedure that allows us to formulate and classify questions in a systematic fashion.

1. Einleitung

Sehr viele Studien untersuchen, welchen Einfluß die Formulierung von Fragen in Fragebogen auf die Antworten der Befragten hat, wie etwa Schuman/Presser (1981), Sudman/Bradburn/Schwarz (1982), Andrews (1984), Molenaar (1986), Alwin/Krosnick (1991), Költringer (1993), Scherpenzeel/Saris (1997). Wenige Studien jedoch sind dem Problem der Übersetzung von Begriffen in Fragen gewidmet (Hox 1997). Bereits in 1968 wies Blalock auf das Problem der sogenannten Operationalisierung hin, als er feststellte, daß es ein Mißverhältnis gäbe zwischen der Sprache der Theorie und jener der empirischen Forschung. Als sich Blalock nach zwei Jahrzehnten wiederum diesem Thema widmete mußte er feststellen, daß dieses Mißverhältnis noch nicht ausgeglichen war.

Im Anschluß an Northrop (1947) unterschied Blalock (1990) zwei Arten von Begriffen: intuitive Konzepte (concepts by intuition) und postulierte Konzepte (concepts by postulation). Intuitive Begriffe beziehen sich zum Beispiel auf kognitive Urteile, Evaluationen, Rechte und Normen und enthalten direkte Anweisungen für den Inhalt der Fragen. Postulierte Begriffe hingegen sind Konstrukte, wie etwa Ethnozentrismus, Macht und Integration und sind aus intuitiven Begriffen aufgebaut. Durch die Operationalisierung von Konstrukten wird angegeben, auf welche intuitiven Begriffe sie sich beziehen. Danach kann anhand der intuitiven Begriffe festgestellt werden, welche Fragen in Betracht kommen. Die Übertragung von Konstrukten in Fragen besteht somit aus zwei Schritten: zunächst müssen Angaben gemacht werden auf welche intuitiven Begriffe sich ein Konstrukt bezieht und danach können für die angegebenen intuitiven Begriffe die geeigneten Fragen formuliert werden. In der sozialwissenschaftlichen Praxis wird jedoch meistens der erste Schritt nicht ausgeführt, so daß direkt anhand des Konstrukts Fragen formuliert werden, wodurch häufig heterogene intuitive Begriffe einem Konstrukt zugeordnet werden.

Einige dieser Probleme wollen wir anhand der Demokratieskala von Kaase (1971) veranschaulichen. Obwohl der Autor dieses Instrument mit großer Sorgfalt entwickelt hatte, weist es dennoch einige Probleme auf, die Blalock erwähnte. Kaases (1971: 142) Demokratieskala stützt sich auf die von Lipset (1963) und Dahrendorf (1965) erstell-

ten Kriterien für ein demokratisch orientiertes politisches System. Diese Kriterien lauten wie folgt:

- a) die Garantie individueller Partizipationsrechte einschließlich des Schutzes von Minderheitsrechten;
- b) die Organisation politischer Herrschaftsinstitutionen nach dem Prinzip der Kontrolle dieser Institutionen durch regelmäßige Wahlen mit der Chance des Auswechslens von Führungspersonal;
- c) die grundsätzliche Anerkennung der Legitimität von Mittel- und Zielkonflikten;
- d) die Ablehnung von Gewalt als Mittel der politischen Auseinandersetzung;
- e) ein Konsens über die fundamentalen demokratischen Wertsetzungen, der überhaupt erst die Voraussetzungen für die regelrechte Austragung der Konflikte schafft.

Anhand dieser Kriterien entwarf Kaase vier Begriffe für demokratische Einstellungen und formulierte Fragen, die sie messen sollten. Tabelle 1 enthält Kaases postulierte Begriffe und Items und die von uns hinzugefügten intuitiven Begriffe, auf die sich die postulierten Begriffe beziehen. Aus der Tabelle ergibt sich eindeutig, daß die Namen der intuitiven Begriffe mit dem Inhalt der Fragen übereinstimmen. Die Beziehung zwischen postulierten und intuitiven Begriffen ist jedoch mehrmals problematisch.

Der erste postulierte Begriff in der Tabelle ist operationalisiert mit zwei verschiedenen, intuitiven Begriffen, d.h. mit einer „kognitiven Bewertung der Opposition“ und einem „Recht zu regieren“. Es ist deshalb auch nicht deutlich, welcher dieser intuitiven Begriffe den postulierten Begriff richtig wiedergibt und weshalb diese intuitiven Begriffe auch noch zusammengefügt werden, um diesen postulierten Begriff anzudeuten. Man könnte auch nur einen Begriff nehmen oder noch andere willkürlich hinzufügen.

Dasselbe gilt für das zweite und das vierte postulierte Konzept. Die Einstellungen zur Legitimität von Konflikten sind angegeben mit einer „Evaluativen Überzeugung von Interessenkonflikten“ und einer „Norm für die Opposition“ und die Einstellungen zu individuellen Grundrechten enthalten neben Rechten auch wiederum eine Norm. Bürklins (1980) Weiterentwicklung der Demokratieskala ist ebenfalls problematisch, zumal er drei neue Items hinzufügt und auch die postulierten Begriffe aus welchen er die Fragen erstellt, anders formuliert. Auch hier ist der Zusammenhang zwischen den postulierten Konzepten und den Fragen nicht deutlich.

Obwohl es nützlich wäre, den Zusammenhang zwischen intuitiven und postulierten Begriffen zu untersuchen, wollen wir uns im Rahmen dieses Aufsatzes jedoch nur mit dem Zusammenhang zwischen einigen intuitiven Begriffen und Fragen der Umfrageforschung befassen.

Tabelle 1: Operationalisierung von Demokratischen Einstellungen

Postulierte Begriffe	Intuitive Begriffe	Items
Einstellungen zur Legitimität von Konflikten	Evaluative Überzeugung von Interessenkonflikten Norm für die Opposition	Die Auseinandersetzung zwischen den verschiedenen Interessengruppen in unserer Gesellschaft und ihre Forderungen an die Regierung schaden dem Allgemeinwohl. Aufgabe der politischen Opposition ist es nicht, die Regierung zu kritisieren, sondern sie in ihrer Arbeit zu unterstützen.
Einstellungen zu politischen Institutionen	Kognitive Bewertung der Opposition Recht zu regieren	Eine lebensfähige Demokratie ist ohne politische Opposition nicht denkbar Jede demokratische Partei sollte grundsätzlich die Chance haben, an die Regierung zu kommen.
Einstellung zur Gewalt als Mittel zur politischen Auseinandersetzung	Evaluative Überzeugung zur Gewaltanwendung	In jeder demokratischen Gesellschaft gibt es bestimmte Konflikte, die mit Gewalt ausgetragen werden müssen.
Einstellungen zu individuellen Grundrechten	Recht auf Demonstration Verlust Recht auf Demonstration Recht auf Meinungsfreiheit Norm für Interessen	Jeder Bürger hat das Recht, notfalls für seine Überzeugung auf die Straße zu gehen. Der Bürger verliert das Recht zu Streiks und Demonstrationen, wenn er damit die öffentliche Ordnung gefährdet Jeder sollte das Recht haben, für seine Meinung einzutreten, auch wenn die Mehrheit anderer Meinung ist. Die Interessen des ganzen Volkes sollten immer über den Sonderinteressen des einzelnen stehen.

Folgende intuitive Begriffe, werden erörtert: kognitive Urteile, Evaluationen, affektive Bewertungen, demographische Merkmale, Werte, kognitive Überzeugungen, Präferenzen, Normen, staatliche Maßnahmen, Rechte, Verhalten, Handlungstendenzen und Erwartungen. Um den Zusammenhang zwischen den intuitiven Konzepten und ihrer sprachlichen Formulierung in den Items oder Behauptungen zu verdeutlichen, werden wir die Strukturen der Aussagen analysieren, die diese Konzepte enthalten.

Da man mit den Begriffen nur bestimmte Fragen stellen kann, werden auch die Frageformen untersucht, die sich für diese Konzepte eignen. Wenn man die Begriffe mit der Frageform verknüpft, erhält man ein Verfahren zur systematischen Formulierung und

Klassifizierung von Fragen. Dieses Instrument kann angewendet werden für die Formulierung von Fragen und auch um Fehler in Fragen zu entdecken, wodurch man eine erhebliche Verbesserung der Qualität der Fragen erzielen kann. Ein derartiges Verfahren wurde von uns bereits für amerikanische und niederländische Umfragen entwickelt (Saris/Gallhofer 1997, 1998).

Aus der praktischen Anwendung ergab sich, daß ein effizientes und zuverlässiges Verfahren benötigt wird, um große Fragebestände schnell analysieren zu können. Sowohl Effizienz als auch Zuverlässigkeit kann man am ehesten mit einem computergestützten Verfahren erreichen. Aus diesem Grund wurde eine computergestützte Kodierungsmethode entwickelt und getestet.

Das Ziel dieses Aufsatzes ist es, diese Methode zu erörtern. Die vorliegende Studie beruht auf deutschsprachigen Fragen, die aus Befragungen des Ifes Institut in Wien stammen und zwischen 1986-1992 ausgeführt wurden.

2. Klassifikation des Frageinhalts

Da wir alle Begriffe anhand von Behauptungen (assertions) definieren, wollen wir zunächst diesen Begriff einführen. Unter einer Behauptung verstehen wir eine Aussage, die Objekte mit Objekten oder Objekte mit Prädikaten verbindet. Behauptungen können somit aus drei Teilen bestehen: Objekten (x,y,...), Connectoren (C) und Prädikaten (p). Die meist allgemeine Form einer Behauptung (a) lautet deshalb (xCy) oder (xCp).

Objekte von Behauptungen können alles enthalten, z.B. leblose Gegenstände (Sachen, Begriffe, politische Maßnahmen) oder Lebewesen (Personen, Gruppen, Institutionen). Objekte werden mit den Symbolen „x, y.“ angedeutet. Für häufig vorkommende Objekte verwenden wir die folgenden Symbole:

- s (subject) bezeichnet den Befragten,
- b (behavior) deutet Verhalten an,
- o (one) weist auf jedermann hin,
- g (government) bezeichnet die Regierung,
- v (value) bedeutet Wert.

Prädikate bestehen aus allerlei Merkmalen, die ein Connector mit einem Objekt verbinden kann. Wir deuten sie mit „p“ an. Häufig enthalten sie Adjektive, wie etwa „groß“, „interessant“, „nett“ und dergleichen. Auch hier geben wir häufig vorkommende Prädikate mit den folgenden Symbolen an:

- e (evaluation) weist auf Evaluationen hin, wie etwa „gut/schlecht“, „wertvoll/wertlos“, „vorteilhaft/unvorteilhaft“,
- f (feeling) deutet affektive Bewertungen an, wie etwa „glücklich/unglücklich“, „angenehm/unangenehm“, „freundlich, unfreundlich“ usw.,
- i (importance) bedeutet „wichtig“,
- n (number) bezeichnet eine Zahl oder Frequenz.

Behauptungen benötigen immer Connectoren. Als allgemeines Symbol, das eine Verbindung andeutet, verwenden wir C (connector). Auch hier haben wir einige spezifische Connectoren definiert:

- I (is) deutet Verbindungen an zwischen Objekten und Prädikaten, die durch die Wörter „ist“ oder „hat“ entstehen.
- F (feeling) weist auf Verbindungen hin, die durch den Gebrauch von Wörtern entstehen, die eine Affektive Bewertung ausdrücken, wie etwa „mögen/nicht mögen“, „sorgen für“ und dergleichen.
- R (relation) bezieht sich auf Relationen wie etwa „Ursache und Wirkung“, „Ziel und Mittel“, „Gleichartigkeit /Gegensätzlichkeit“, „Abstand/Nähe“ und dergleichen.
- H (has to) deutet Verbindungen an, die durch Wörter wie etwa „soll“ und „muß“ realisiert werden.

Mit den soeben angeführten Symbolen, können wir nun die Begriffe beschreiben, die sich auf Behauptungen in Fragen aus Umfragen beziehen.

Kognitive Überzeugungen (cognitive beliefs) gelten in der psychologischen Literatur als eine Basiskomponente des Attitüde Begriffs (Kretch/Crutchfield 1948; Bradburn/ Sudman 1988; Fishbein/Ajzen 1975; Ajzen 1989; Eagly/Chaiken 1993; Van der Pligt/ De Vries 1995). Zwei Sorten kognitiver Begriffe werden in der Literatur unterschieden. Der eine wird Perzeption oder Urteil genannt und der andere heißt Relation oder Beziehung. Die Struktur einer Behauptung (assertion), die sich auf ein kognitives Urteil (judgment) bezieht (a_i), kann wiedergegeben werden mit (xIc) , was bedeutet „x hat das Merkmal c“.

Die Struktur einer Behauptung, die eine Relation enthält (a_r), kann wiedergegeben werden mit (xRy) , was bedeutet „x hat eine Beziehung zu y“. Diese Behauptung kann andeuten, daß x die Ursache von y ist, oder etwa, daß x Einfluß auf y hat, und dergleichen. In der Attributionstheorie (Kelley/Michela 1980) befaßt man sich unter anderem mit Kausalperzeptionen. Andere Relationen, die häufig Gegenstand sozialwissenschaftlicher Forschung sind, beziehen sich auf Gleichartigkeit/Gegensätzlichkeit von Objekten, Abstand/Nähe zwischen Objekten (Rabinowitz/Macdonald/Lishuag 1991)

oder auf die Mitgliedschaft in einer Organisation und den Zusammenhang zwischen Objekten (Harary 1971; Helmers/Mokken/Plijter/Stokman 1975; Knoke/Kuklinski 1982).

Evaluationen (evaluations) werden meistens in der Literatur (Fishbein/Ajzen 1975; Bradburn/Sudman 1988; Zanna/Rempel 1988; Van der Pligt/De Vries 1995; Tesser/Martin 1996) als eine beobachtbare Komponente von Attitüden gesehen. Die Struktur von evaluativen Behauptungen (a_e) ist (xI_e) und besagt, daß x gut oder schlecht ist. Wir weisen darauf hin, daß die Struktur von evaluativen Behauptungen die gleiche ist wie die von kognitiven Urteilen. Die Begriffe unterscheiden sich jedoch bezüglich des Prädikats. Evaluationen enthalten Bewertungen im Sinn von „gut/schlecht“ oder „utilitär nützlich/unnützlich“, während das Prädikat bei kognitiven Urteilen alles andere enthalten kann, außer evaluativen und emotionalen Prädikaten. Da es auch andere evaluative Strukturen gibt mit positiver oder negativer Konnotation, werden diese im nächsten Abschnitt besprochen.

Unter *evaluativen Überzeugungen* (evaluative beliefs) verstehen wir im Anschluß an Oskamp (1991) Behauptungen über Verhalten oder Relationen, wobei die Verhaltens- oder Relationskomponente eine Bewertung in der Bedeutung von „gut“ oder „schlecht“ enthält. Die Struktur einer evaluativen Überzeugung mit einer Relationskomponente kann angedeutet werden mit $a_{eb}=xR_e y$, was etwa besagen könnte, „ x übt einen günstigen Einfluß auf y aus“. Wenn die Struktur eine evaluative Verhaltenskomponente enthält, wird sie angegeben mit $a_{eb}=xD_e y$, was bedeuten könnte, „ x bedroht y “. Deutlichkeitshalber sei darauf hingewiesen, daß auch andere Behauptungen eine Evaluationskomponente haben können, so daß evaluative Überzeugungen die Struktur von verschiedenen Behauptungen annehmen können.

Affektive Bewertungen (feelings) wurden in der Literatur als Evaluationen betrachtet (Bradburn/Sudman 1988). Seit einigen Jahren jedoch trennen einige Umfrageforscher kognitive Evaluationen von affektiven Bewertungen (Bagozzi 1989; Ajzen 1991). In unserer Klassifikation definieren wir drei Elementarbehauptungen, die sich auf affektive Bewertungen beziehen. Die erste Behauptung (a_f) lautet (sFx) und kann zum Beispiel andeuten, „ x liebt y “. Die zweite Struktur ist (xIf) und kann etwa besagen, „ x ist angenehm“. Die dritte Struktur (xRf) kann andeuten, „ x stimmt mich traurig“. Zusammenfassend kann gesagt werden, daß „ f “ oder „ F “ Emotionen andeuten, wie etwa „Angst, Abscheu, Ärger, Kummer, Geringschätzung, Scham, Erniedrigung, Leidenschaft, Glück, Überraschung“ und dergleichen (Cornelius 1996).

Nach *Präferenzen* (preferences) wird häufig in Erhebungen gefragt, die sich mit Konsumenten- und Entscheidungsverhalten befassen. Öfters wird ein Vergleich gemacht

zwischen dem attraktivsten Objekt und den weniger begehrten Objekten (Torgerson 1958; Von Winterfeld/Edwards 1986). Die Struktur eines derartigen Vergleichs (a_{pr}) ist ($xPy, z...$) und lautet, „der Befragte zieht x allen anderen Objekten vor“. In diesem Fall findet ein Vergleich zwischen allerlei Objekten statt, wobei ein Objekt als das bevorzugte angewiesen wird. Man könnte die Präferenz auch ausdrücken mit „x ist das beste Objekt“ oder mit „x wird bevorzugt“. Wir haben jedoch die Struktur ($xPy, z...$) gewählt, um deutlich zu machen, daß eine Auswahl aus mehreren Objekten stattfindet. Eine andere Weise, Präferenzen zu formulieren besteht darin, daß man den Befragten vor die Wahl stellt, ob er „dafür“ oder „dagegen“ ist. Hier muß man also nur eine Auswahl zwischen zwei Objekten treffen. Diese zweite Vorzugsbehauptung kann mit (sPy) wiedergegeben werden, was bedeutet, daß der Befragte (s) einen Standpunkt (y) billigt oder nicht billigt.

Werte (values) sind der nächste Begriff den wir besprechen wollen. Die Struktur einer Behauptung, die einen Wert enthält (a_i), formulieren wir als (vI), was bedeutet, „v ist wichtig“. Diese Behauptung hat die gleiche Form wie kognitive Urteile, Evaluationen und die zweite Struktur von affektiven Bewertungen. Der Unterschied besteht jedoch im Prädikat, das in diesem Fall „wichtig“ ist und in der Andeutung eines Wertes „v“, worunter man zentrale Lebensziele versteht, wie etwa „Ehrlichkeit, Sicherheit, Gerechtigkeit, Glück“ und dergleichen (Rokeach 1973).

Normen (norms) spielen eine wichtige Rolle in der sozialwissenschaftlichen Forschung (Sorokin 1928; Parsons 1951; Homans 1965). Coleman (1990: 242) definiert Normen als eine Aufstellung von Verhaltensweisen, „welche von einer Gruppe von Personen als richtig und zutreffend betrachtet wird“. Normen beziehen sich also auf individuelle Wahrnehmungen von sozial und kulturell gebilligten Verhaltensregeln, das heißt, wie jemand sich in einer bestimmten Situation verhalten soll. Die Struktur dieser Behauptung (a_n) ist (oHb), was besagt, „jemand soll b ausführen“.

Staatliche Maßnahmen (policies) werden häufig untersucht, um festzustellen, wie Bürger über Pläne oder Handlungsweisen der Behörden denken (Sniderman/Brody/Tetlock 1991; Holsti 1996). Eine staatliche Maßnahme (a_p) hat die Struktur (gHb) und besagt, „die Regierung soll b ausführen“. Der Unterschied zu einer Norm besteht darin, daß der Ausführende kein Individuum ist sondern eine Behörde.

Rechte (rights) oder Fragen nach bürgerrechtlichen Angelegenheiten werden ebenfalls öfters gestellt (Sniderman/Brody/Tetlock 1991). Rechte bestehen aus individuellen Wahrnehmungen, die angeben, wer berechtigt ist, etwas zu tun. Die Struktur einer Rechtsbehauptung (a_r) ist ($oIRy$) und lautet „x darf y ausführen“.

Handlungstendenzen (action tendencies) werden häufig beschrieben als die dritte Komponente des Attitüde Begriffs (Ajzen/Fishbein 1980; Bradburn/Sudman 1988; Sudman/Bradburn/Schwarz 1982; Eagly/Chaiken 1993). Eine Handlungstendenz richtet sich auf zukünftiges, zu erwartendes Verhalten. Man nimmt dabei an, daß die zukünftige Verhaltenstendenz, eine gute Vorhersage für das tatsächliche Verhalten ist. Eine Verhaltenstendenzbehauptung (a_i) hat die Struktur (xFDy) und besagt, „x wird y ausführen“.

Erwartungen zukünftiger Ereignisse (expectations of future events) kommen in Umfragen ebenfalls häufig vor (Graesser/Bommareddy/Swamer/Golding 1996). Diese Behauptungen befassen sich mit zukünftigen Ereignissen, auf die man selbst keinen Einfluß hat. Die Struktur dieser Behauptung (a_{ex}) ist (xEy) und lautet „x erwartet y“.

Verhalten (behavior) besteht aus Handlungen oder Aktivitäten, die sich in der Gegenwart ereignen oder in der Vergangenheit stattgefunden haben (Sudman/Bradburn/Schwarz 1982; Smith 1987). Die Struktur einer Verhaltensaussage (a_b) ist (sDy). Ausdrücke wie etwa „x führt y aus“ oder „x hat y unternommen“ werden dieser Aussage zugeordnet. Die Struktur ist die gleiche wie die von Handlungstendenzen. Inhaltlich besteht jedoch ein großer Unterschied zwischen den Behauptungen: Handlungstendenzen sind subjektiv und beziehen sich auf die Zukunft. Verhalten dagegen ist faktisch und kann im Prinzip kontrolliert werden, da es sich auf die Vergangenheit oder Gegenwart bezieht.

Demographische Merkmale (demographic variables) werden sehr häufig in Befragungen erhoben und alle bestehenden Klassifikationen erwähnen diese Merkmale (Oppenheim 1966; Sudman/Bradburn/Schwarz 1982; Converse/Schuman 1984; Smith 1987; Bradburn/Sudman 1988). Es handelt sich um Variablen wie Alter, Beruf, Geschlecht, Religion und dergleichen. Die Struktur von demographischen Behauptungen (a_d) ist die gleiche wie die von kognitiven Urteilen. Die Struktur lautet (xIc), wobei x der Befragte ist und c ist z. B. der Beruf .

Somit gibt es sieben verschiedene Behauptungsstrukturen, welche die meisten in Befragungen verwendeten intuitiven Begriffe beschreiben. Tabelle 2 faßt sie zusammen. Innerhalb jeder Klasse wird anhand der Bedingungen angegeben, ob ein bestimmter Begriff verfügbar ist oder nicht.

Tabelle 2 : Zusammenfassung der Strukturen der elementaren* Behauptungen

Strukturen	Bedingungen	Intuitiver Begriff
Beschreibungen		

(xIp)	wenn x ist nicht v und wenn p ist nicht e, i, f oder d und wenn p ist e und wenn p ist f und wenn p ist d wenn x ist v und wenn p ist i	Kognitives Urteil Evaluation Affektive Bewertung Demographisches Merkmal Wert
Relationen (xRy)	wenn R keine Konnotationen enthält wenn R eine evaluative Konnotation enthält wenn R eine affektive Konnotation enthält	Kognitive Überzeugung Evaluative Überzeugung Affektive Bewertung
Präferenzen (xPy, z...)	wenn man aus mehreren Objekten wählen kann	Präferenz
(xPy)	wenn man nur aus zwei Objekten wählen kann	Präferenz
Pflichten (xHy)	wenn x ist o wenn x ist g	Norm Staatliche Maßnahme
Rechte (xIRy)	normalerweise	Recht
Handlungen (xDy)	normalerweise wenn D mit F substituiert wird wenn D mit FD substituiert wird wenn D eine evaluative Konnotation hat	Verhalten Affektive Bewertung Handlungstendenz Evaluative Überzeugung
Erwartungen zukünftiger Ergebnisse (xEy)	normalerweise	Erwartung

*Die Formulierung von zusammengesetzten Behauptungen erfolgt durch die Substitution von x oder y durch eine neue Behauptung (a) oder durch die Zufügung eines Nebensatzes zu einem Gegenstandsausdruck.

Tabelle 2 eignet sich auch für die Frageformulierung. Wenn man Behauptungen für einen spezifischen Begriff formulieren will, geben die Bedingungen in Tabelle 2 an, wie dieser intuitive Begriff auszudrücken ist. Nehmen wir an, wir wollen eine Evaluation des Begriffs Asylsuchender formulieren. Die Tabelle gibt als Struktur (xIe) an, was sich etwa in die folgende Behauptung umsetzen läßt: „Asylsuchende sind nützlich als Hilfsarbeiter“. Wenn man eine affektive Bewertung ausdrücken will, verfügt man über die Struktur (xIf) und man könnte zum Beispiel die folgende Behauptung formulieren: „Asylanten sind im allgemeinen freundliche Menschen“. Ein kognitives Urteil (xIc) wäre etwa: „Asylsuchende sind Leute aus der dritten Welt“. Eine Relation (xRy) könnte

man wiedergeben mit dem nächsten Beispiel: „Da diese Leute in ihrem eigenen Land politische oder wirtschaftliche Probleme haben, kommen sie zu uns“. Auf gleiche Weise lassen sich Behauptungen für andere Begriffe formulieren.

Außer den elementaren Behauptungen gibt es auch zusammengesetzte, d.h., man kann aus Behauptungen noch andere ableiten. Einige häufiger vorkommende abgeleitete Behauptungen seien noch kurz erwähnt.

Wichtigkeit eines Urteils (importance of a judgment) (a_{ij}) gibt wieder, inwieweit ein Befragter eine Behauptung persönlich als wichtig empfindet. Krosnick/Abelson (1991) betrachten „Wichtigkeit“ als eine Dimension von „Attitudenstärke“. Wir verwenden diesen Begriff hier jedoch im weiteren Sinn. Die Struktur der Wichtigkeit eines Urteils ist (a_i) und besagt, „Behauptung a ist wichtig“. Man kann grundsätzlich bei allen Behauptungen nach ihrer Wichtigkeit fragen.

Sicherheit eines Urteils (certainty of a judgment) (a_{ce}) zeigt an, inwieweit ein Befragter vom Inhalt einer Behauptung überzeugt ist. Krosnick/Abelson (1991) erwähnen auch diesen Begriff im Zusammenhang mit Attitüden. Die Struktur dieser Behauptung ist (a_{lce}) und lautet „ich bin mir sicher, daß Behauptung a so ist“. Auch hier gilt, daß man grundsätzlich bei allen Behauptungen nach dem Überzeugungsgrad des Befragten fragen kann.

Ermittlung der Intensität einer Behauptung (appraisal of an assertion) (a_a) befaßt sich mit der Stärke einer Relation oder eines Verhaltens. Die Struktur $((xRy) \text{ I } c)$ kann besagen „die Relation zwischen x und y ist sehr stark“ und die Struktur $((xDy) \text{ I } c)$ kann bedeuten, „Verhalten a ist sehr zweckmäßig“.

Krosnick und Abelson (1991) verwenden diesen Begriff im Zusammenhang mit affektiven Bewertungen hinsichtlich eines Attitüdeobjektes. Wir definieren diesen Begriff jedoch in einem weiteren Sinn, nämlich in bezug auf Relationen und Verhalten.

Häufigkeit von Verhalten (frequency of behavior) (a_{bn}) hat die Struktur $((xDy) \text{ I } n)$ und bedeutet, „x macht oft y“.

In dieser Übersicht haben wir die meisten Begriffe beschrieben, die in der Umfrageforschung verwendet werden. Mit dieser ausführlichen Beschreibung beabsichtigten wir, die Struktur möglicher Behauptungen zu erklären.

3. Klassifikation von Frageformen

Außer Grundtypen von Behauptungen gibt es auch noch Grundtypen von Frageformen. Frageformen werden in verschiedenen Wissenschaftszweigen erforscht, wie etwa

in der Linguistik, Philosophie und in den Sozialwissenschaften (Kearsley 1976; Dillon 1984; Meyer 1988; Graesser/Bommareddy/Swamer/Golding 1996). Alle diese Studien betonen die Wichtigkeit der sogenannten „W-Fragewörter“ (wh-interrogatives): „wer, was, welche, wozu, wo, wann, warum, wie“. Linguisten und Philosophen unterscheiden auch eine Frageform, welche die Antwort „ja“ oder „nein“ bewirken soll. Linguistisch betrachtet besteht diese Frageform aus der Umkehrung der Wortfolge. Das Verb steht nämlich an erster Stelle. Beispiele hierzu sind etwa, „Denken Sie, daß diese Maßnahme der Europäischen Union günstig für unser Land ist?“ oder „Ist das Ihr Lebensziel?“

Sozialwissenschaftler haben noch andere Frageformen hinzugefügt, wie etwa Fragen, die eine Zustimmung oder Ablehnung erfordern. Anhand dieser Studien und einer Analyse von Umfrage-Fragen beschreiben wir nun die relevanten Frageformen:

Einfache Frage (simple Q): $Q(a)$ =Ist a wahr oder nicht wahr? Beispiel: $Q(xIc)$ Ist Ihrer Meinung nach Berlin eine schöne Stadt?

Diese Frageform wird sehr häufig verwendet in der Umfrageforschung. Sie ermittelt an welcher Seite der Befragte steht, d.h. sie gibt die Richtung der Meinung an.

Ausführlichere Frage (extended Q): $HQ(a)$ =Wieviel von c hat x? Beispiel: $HQ(xIc)$ Finden Sie, daß Berlin sehr schön, ziemlich schön, ziemlich häßlich oder sehr häßlich ist? Diese Frageform mißt die Gradierung einer Eigenschaft bei einem Objekt. Krosnick/Abelson (1991) nennen diesen Begriff „Extremität“.

Zustimmung (agreement A): $A(a)$ =Stimmen Sie dieser Behauptung zu oder lehnen Sie sie ab? Beispiel: $A(a)$: Stimmen Sie der folgenden Aussage zu oder lehnen Sie sie ab: „Es gibt einen Gott im Sinne eines höheren Wesen“?

Diese Frageform mißt die gleiche Information wie die einfache Frage, da die Zustimmung zur Behauptung ebenfalls angibt, auf welcher Seite der Befragte steht.

Ausführlichere Zustimmung (extended agreement HA): $HA(a)$ =Sagen Sie mir bitte, ob Sie mit dieser Aussage voll übereinstimmen, eher übereinstimmen, eher nicht übereinstimmen oder gar nicht übereinstimmen? Beispiel: $HA(a)$: Wie sehr stimmen Sie mit dieser Aussage überein: „Es gibt einen Gott im Sinne eines höheren Wesens?“

Diese Frageform mißt nicht den Grad des Vorhandenseins einer Eigenschaft, sondern das Ausmaß der Zustimmung. Krosnick/Abelson (1991) nennen diese Frageform Intensität der Zustimmung zu einer Behauptung.

Was (WHAT): $What(a)$ =Was halten Sie von x? Beispiele: $What(sDy)$: „Was haben Sie dort gekauft?“ $What(xRy)$: „Was war die Ursache Ihres Verkehrsunfalls?“ $K(sDy)$: „Was wissen Sie von der Steuerreform?“

Diese Frageform eignet sich für verschiedene Frageinhalte, wie etwa Faktfragen, kognitive Urteile und Wissen.

Wer (WHO): Who(a)=Wer ist...? Beispiel: Who(xIc): „Wer ist der Generalsekretär der Vereinten Nationen?“

Mit dieser Frage will man Informationen über Personen erhalten. Meistens werden diese Fragen angewandt, um das Wissen des Befragten hinsichtlich einer Person zu ermitteln.

Wann (Zeit) (When: time): T(a)=Wann hat x y unternommen? Beispiel: T(xDy): „Wann haben Sie zum ersten Mal die Schule besucht?“

Mit dieser Frage will man den Zeitpunkt ermitteln, an dem eine Handlung stattgefunden hat.

Wo (Ort) (Where: place): P(a)=Wo hat x y unternommen? Beispiel: P(xDy): „Wo haben Sie zum ersten Mal die Schule besucht?“

Es ist deutlich, daß diese Frage den Ort ermittelt an, dem ein Ereignis stattgefunden hat.

Wie (Verfahren) (How: procedure): HPR(a)=Wie hat x y ausgeführt? Beispiel: HPR(xDy): Wie sind Sie dorthin gefahren?

Mit dieser Frageform will man die Art und Weise ermitteln, wie ein Verhalten stattgefunden hat.

Warum (Why): WHY(a)=Warum hat x y gemacht? Beispiel: WHY(sDy): Warum hat er ein Auto gekauft?

Diese Frage befaßt sich mit den Gründen oder Ursachen eines Verhaltens.

Welche/Was für ein (Which): WHICH(a)=Welches x ist c? Beispiele: WHICH(sDy): Was für ein Buch lesen Sie derzeit? W(xDy): Welche Sicherheitsmaßnahmen hat Ihr Betrieb getroffen? W(xPy,z.): Welche Wohnung ziehen Sie vor?

Diese Frageform wird häufig für Präferenzen verwendet, um anzugeben, welchen Gegenstand man bevorzugt.

Wissen (Knowledge: K): K(a)=Kennt x y? Beispiel: K(xIc): Wissen Sie, wer der Generalsekretär der Vereinten Nationen ist?

Diese Frageform wird häufig angewandt, um nach der Kenntnis der Befragten auf verschiedenen Gebieten zu informieren. Wie bereits zuvor erwähnt, erfüllen die „wer“ und „was“ Fragen ebenfalls diesen Zweck.

Hiermit haben wir die Frageformen, die in der Umfrageforschung verwendet werden behandelt. Tabelle 3 faßt die Strukturen der Frageformen zusammen.

Aus dem Vorgegangenen ist ersichtlich, daß wir davon abgesehen haben, die Antwortkategorien anzuführen, insofern sie nicht bereits in der Frageformulierung erwähnt waren. Der Grund dafür ist, daß wir uns hauptsächlich mit der Frageformulierung befassen wollten, zumal Antwortalternativen bereits in vielen Studien (vgl. Schuman/Presser 1981; Andrews 1984; Költringer 1993; Scherpenzeel/Saris 1997) eingehend untersucht wurden.

Tabelle 3: Zusammenfassung der Strukturen der Frageformen

Frageform	Struktur	Ermittelte Information
Einfache Frage: Q	xIp?	Richtung der Meinung
Ausführlichere Frage: HQ	Wieviel von eIx?	Extremität einer Eigenschaft
Zustimmung: A	Stimmen Sie (a) zu ?	Zustimmung
Ausführlichere Zustimmung: HA	Wie sehr stimmen Sie (a) zu ?	Intensität der Zustimmung
Was: WHAT	Was (a) ?	Objekte
Wer: WHO	Wer (a) ?	Personen
Wann: T	Wann (a) ?	Zeitpunkt
Wo: P	Wo (a) ?	Ort
Wie: HOW	Wie (a) ?	Verfahren
Warum: WHY	Warum (a) ?	Gründe
Welche, was für ein: WHICH	Welche (a) ?	Welches von einer Menge Objekten
Wissen: K	Kennen Sie (a) ?	Wissen

4. Frageformulierung

Wenn man den Frageinhalt mit der Frageform verknüpft, entsteht eine große Anzahl möglicher Fragen. Tabellen 4 und 5 veranschaulichen diese Möglichkeiten. Ein „+“ Zeichen in den Tabellen weist darauf hin, daß eine Frage in einer bestimmten Kombination denkbar ist und ein „-“ Zeichen deutet an, daß diese Frage nicht gangbar ist. Aus den Tabellen ist ersichtlich, daß Behauptungen mit derselben Grundstruktur auch meistens mit denselben Frageformen kombiniert werden können. Um die Anwendung der Klassifikation zu veranschaulichen, geben wir anhand von einigen Beispielen wieder, wie man ein Urteil über die Regierung erfragen kann. Unser Hauptanliegen ist es die Vielfalt der Frageformen aufzuzeigen und wir lassen daher denkbare Abwechslungen der Antwortkategorien und Instruktionen außer Betracht. Die Symbole die bei den Beispielen angeführt sind, beziehen sich auf Tabelle 4. Wir kombinieren jeweils einen Frageinhalt mit einer Frageform, die eine bestimmte Zelle in der Tabelle darstellen.

Tabelle 4: Entwicklung von Fragen durch Kombination von Frageinhalten mit Frageformen

Intuitive Begriffe	Q	HQ	A	HA	WHAT	Frageformen WHO	WHICH	T	P	HPR	WHY	K
Beschreibungen												
Kognitive												
Urteile $a_i = (xIc)$	+	+	+	+	+	+	+	-	-	-	+	+
Evaluationen $a_e = (xIc)$	+	+	+	+	+	+	+	-	-	-	+	-
Werte $a_v = (vIi)$	+	+	+	+	+	+	+	-	-	-	+	-
Relationen $a_r = (xRy)$	+	+	+	+	+	+	+	-	-	-	+	+
Evaluative Überzeugungen												
$a_{ib} = (xDey)$	+	+	+	+	-	-	-	-	-	+	+	-
$a_{eb} = (xRey)$	+	+	+	+	-	-	-	-	-	+	+	-
Affektive Bewertungen												
$a_{if} = (xIf)$	+	+	+	+	+	+	+	+	+	-	+	-
$a_{if} = (sFy)$	+	+	+	+	+	+	+	+	+	-	+	-
$a_{if} = (xRf)$	+	+	+	+	+	+	+	+	+	-	+	-
Präferenzen												
$a_{pr} = (xPy, z, \dots)$	+	+	+	+	+	-	-	-	-	-	+	-
$a_{pr} = (sPy)$	+	+	+	+	+	-	-	-	-	-	+	-
Rechte und Pflichten												
Nomen $a_n = (oHb)$	+	+	+	+	+	+	+	+	+	-	+	+
Staatliche Maßnahmen $a_p = (gHb)$	+	+	+	+	+	+	+	+	+	-	+	+
Rechte $a_r = (oIRb)$	+	+	+	+	+	+	+	+	+	-	+	+
Verhalten												
Handlungstendenzen $a_t = (xFDy)$	+	+	+	+	+	+	+	+	+	+	+	+
Verhalten $a_b = (xDy)$	+	+	+	+	+	+	+	+	+	+	+	+
Erwartungen $a_{ex} = (xEy)$	+	+	+	+	+	+	+	+	+	-	+	+
Demograph. Merkmale $a_d = (xIc)$	+	-	-	+	+	+	+	+	+	-	-	-

Tabelle 5: Entwicklung von Fragen durch Kombination von Frageinhalten mit Frageformen

Intuitive Begriffe	Q	HQ	A	HA	WHAT	Frageformen WHO	WHICH	T	P	HPR	WHY	K
Wichtigkeit eines Urteils $a_i = (aI)$	-	+	-	-	-	-	-	-	-	-	+	-
Sicherheit eines Urteils $a_c = (aIcc)$	-	+	-	-	-	-	-	-	-	-	+	-
Intensität eines Urteils $a_s = ((sRy) Ic)$ $a_n = ((sDy) Ic)$	+	+	-	-	-	-	-	-	-	-	-	-
Häufigkeit von Verhalten $a_{bn} = ((sDy) I n)$	-	+	-	-	-	-	-	-	-	-	-	-

Die einfachste Art und Weise um nach einem Urteil zu fragen, ist wie folgt:

$W(a_i)$: *Was halten Sie von unserer Regierung?*

Dies ist eine offene Frage, wobei die Befragten selbst ein Prädikat wählen und es bewerten können. Verschiedene Befragte können dabei an unterschiedliche Prädikate denken, wie etwa an Konsequenzen von Regierungsbeschlüssen oder an die Zusammenarbeit innerhalb der Regierung oder an Regierungserklärungen und dergleichen. Da diese Prädikate nicht vergleichbar sind, ist es oft besser, eine geschlossene Frage zu stellen, in der das Prädikat explizit angegeben ist, so wie zum Beispiel:

$HQ(a_c)$: *Wie gut finden Sie unsere Regierung: 1 sehr gut, 2 gut, 3 weder gut noch schlecht, 4 schlecht, 5 sehr schlecht.*

Hier ist deutlich angegeben, daß die Befragten die Qualität der Regierung beurteilen sollen, wobei sie über fünf Antwortalternativen verfügen.

Wenn man eine Q-Frage stellt, werden die Antwortalternativen ziemlich eingeschränkt:

$Q(a_s)$: *Finden sie unsere Regierung eher gut oder eher schlecht?*

Bei dieser Frage verfügen die Befragten über zwei Antwortmöglichkeiten, was für manche der Befragten zu wenig ist und mehr Antwortverweigerungen zur Folge haben kann (Klingemann 1997).

Man kann auch Behauptungen anwenden, wie etwa:

A(a): Sagen Sie mir bitte zu der folgenden Aussage, ob Sie ihr zustimmen oder sie ablehnen: „Ich halte unsere Regierung für gut.“

Die Antwortalternativen sind noch immer ziemlich eingeschränkt, aber nun wird nach einer Zustimmung zu einer Behauptung gefragt. Auch kann angenommen werden, daß nicht alle Befragten, welche die Regierung positiv beurteilen, der Behauptung zustimmen, da diejenigen, welche die Regierung als besonders gut beurteilen, die Wertung in der Behauptung zu gering finden und folglich mit „nicht zustimmen“ antworten könnten.

Die Wahl der Behauptung ist jedoch problematisch, da sie ziemlich willkürlich ist. Wir geben hier einige Beispiele von alternativen Behauptungen:

*„Ich halte unsere Regierung für sehr gut“
 „Ich halte unsere Regierung für schlecht“
 „Ich halte unsere Regierung für sehr schlecht“*

Mit Sicherheit ist zu erwarten, daß die Wahl der Behauptung die Antwortverteilung beeinflusst (Molenaar 1986). Es ist jedoch nicht deutlich, ob diese Verteilungen vergleichbar sind und ob die Relationen mit anderen Merkmalen gleichbleiben, ungeachtet der Wahl der Behauptung. Saris (1998) fand in einer Studie große Unterschiede in Antworten, die durch die Position der Behauptungen auf der Skala verursacht wurden.

Zum Abschluß führen wir noch ein Beispiel einer HA-Frage an:

HA(a): Sagen Sie mir bitte zu der folgenden Behauptung, ob Sie mit ihr voll übereinstimmen, eher übereinstimmen, eher nicht übereinstimmen oder gar nicht übereinstimmen? „Ich halte unsere Regierung für sehr gut“.

Auch hier kann man die Behauptung variieren, was dieselben Konsequenzen hat, wie zuvor erwähnt. Außerdem behaupten Krosnick/Abelson (1991) mit Recht, daß man mit dieser Frage nicht die Extremität einer Meinung mißt, sondern die Intensität womit man einer Meinung zustimmt.

Wir hoffen, deutlich gemacht zu haben, daß die verschiedenen Fragetypen (Inhalt und Form) nicht willkürlich verwendet werden können. Es hängt von der Formulierung der Frage ab, welche Information man vom Befragten erwarten kann. Aus diesem Grund muß als nächstes geklärt werden, welche Frageformulierung sich am besten für die verschiedenen Begriffe eignet.

Die Klassifikation (Tabellen 4 und 5) gibt deutlich an, welche Fragestrukturen sich für bestimmte intuitive Begriffe eignen. Sie kann deshalb sehr nützlich sein, wenn man postulierte Begriffe mit übereinstimmenden, intuitiven Begriffen operationalisieren will. Sie kann ebenfalls verwendet werden, um bei bereits vorhandenen Skalen zu überprüfen, ob die Fragen, die einen spezifischen postulierten Begriff messen, homogen sind. Zudem eignet sie sich für das Auffinden von Fehlern in bereits vorhandenen Fragebögen.

Diese Zielsetzungen erfordern ein zuverlässiges Klassifikationsverfahren, das auch effizient angewendet werden kann. Aus diesen Gründen wurde eine computergestützte Methode bevorzugt. Die Entwicklung dieses Verfahrens ist deshalb das weitere Thema dieses Aufsatzes.

5. Die Entwicklung eines zuverlässigen Klassifikationsverfahrens

Für die Entwicklung dieses Verfahrens hatten wir elf Fragebögen des Ifes Instituts in Wien zur Verfügung, die zwischen 1986 und 1992 erhoben wurden. Diese Umfragen befassen sich mit sehr verschiedenen Themen, wie etwa Wahlprognosen, Konsumentenverhalten und Studien der Lebensqualität. Obwohl man nicht von einer repräsentativen Stichprobe sprechen kann, waren die Fragen qua Inhalt und Form sehr verschiedenartig, so daß sie sich für unsere Zwecke eigneten.

Die Fragen in den uns zur Verfügung stehenden Fragebögen wurden zunächst nach Frageinhalten und Frageformen handkodiert. Auf diese Weise erhielten wir 640 kodierte Fragen. Danach entwickelten wir ein Verfahren, um diese kodierte Kategorien mit dem Computer zu reproduzieren. Dies war mit relativ einfachen Mitteln möglich. Mit dieser Methode kann man kontrollieren, inwieweit das halbautomatische Computerverfahren die gleichen Kodierungen ergibt wie die Handkodierungen. Da jedoch zunächst ein Computerverfahren entwickelt werden mußte, um die Frageinhalte und Frageformen zu finden und danach ein Test ausgeführt werden mußte, der ausweisen sollte, ob sich die Klassifizierung tatsächlich für deutsche Fragen eignet, haben wir unsere Daten halbiert. Anhand der ungerade nummerierten Fragen wurde das Computerverfahren entwickelt und mit den gerade nummerierten Fragen wurde der Test ausgeführt. Im folgenden wird das computergestützte Verfahren beschrieben. Zunächst befassen wir uns mit dem Entwurf des Verfahrens, sowohl für die Frageinhalte als auch die Frageformen. Danach werden die Testergebnisse mit der anderen Hälfte der Fragen behandelt und zuletzt wird die Übereinstimmung der Handkodierungen mit den Computerkodierungen kurz erörtert.

5.1 Ein computergestütztes Verfahren zur Klassifikation von Fragen

Wir nennen das Verfahren halbautomatisch, da es derzeit mit der verfügbaren Software noch nicht möglich ist, das Verfahren vollkommen zu automatisieren. Dies ist zum Teil zurückzuführen auf die Komplexität der Sprache. Da es sehr viele Möglichkeiten gibt, um ein und denselben Begriff zu formulieren, ist nicht zu erwarten, daß man bereits in einem ersten Ansatz alle möglichen Formulierungen für einen Frageinhalt oder eine Frageform gefunden hat, um Kriterien dafür aufzustellen. Wir erwarten, daß erst nach der Analyse einer sehr großen Anzahl von Fragebögen die notwendigen Wortkriterien festgelegt werden können. Derzeit wird man noch häufig im Vorhinein manuell mehrdeutige Wörter auffinden müssen, um die für diese Studie relevante Bedeutung zu ermitteln. Dies trifft auch für evaluative Wörter und einige andere inhaltliche Begriffe zu. Für Frageformen trifft das weniger zu, da sie mehr aus Strukturwörtern bestehen. Das entwickelte Verfahren ist sequentiell, was besagt, daß sich die Kodierung der Fragen in einer bestimmten Abfolge vollzieht. Diese Methode bietet den Vorteil, daß man damit schwierige Probleme einfacher lösen kann.

5.2 Die Entwicklung der computergestützten Kodierungsmethode für Frageinhalte

Anhand der Ergebnisse, die wir mit niederländischen Fragen erzielt hatten (Saris/Gallhofer 1998), wurde dasselbe sequentielle Verfahren auch hier angewandt.

1. Zuerst wurde nach den charakteristischen Merkmalen von evaluativen Überzeugungen (a_{cb}) gesucht. Es erwies sich als sinnvoll, mit diesem Frageinhalt zu beginnen, da diese Behauptungen sowohl Merkmale von Evaluationen (a_c) als auch von Relationen (a_r) enthalten. Wenn man zuerst nach Evaluationen oder Relationen suchen würde, würde man auch häufig evaluative Überzeugungen finden. Aus diesem Grund ist es effizienter, zuerst nach a_{cb} zu suchen mit Wortkombinationen, die sowohl a_c als auch a_r enthalten (siehe 2 und 3). Es gibt allerdings auch vereinzelte Wörter und Redewendungen, die eine evaluativ gefärbte Relation oder ein evaluativ gefärbtes Verhalten in sich haben, wie etwa „beschmutzen, verbessern, sorgen, mißbrauchen, die Interessen berücksichtigen, unterstützen...“.

2. Danach wurden die evaluativen Fragen mit Bewertungen im Sinn von „gut/schlecht“ und „utilitär nützlich/unnützlich“ gesucht. Beispiele hierfür sind: „gut, besser, schlecht, schlechter, positiv, günstig, zahlt sich aus, falsch, zu wenig, erfolgreich, eine neue Chance, Vorteil, Nachteil, und dergleichen.“

3. An dritter Stelle begaben wir uns auf die Suche nach Behauptungen, die Relationen (a_r) enthalten. Wörter wie etwa „Gründe, Hauptgrund, Anlaß, beeinflussen, wozu, Voraussetzung, weit entfernt, teilnehmen an...“, und dergleichen“ beziehen sich auf a_r .

4. In der vierten Phase versuchten wir Urteile aufzufinden. Zunächst wurden Werte (a_i) gesucht. Wörter welche die „Wichtigkeit“ ausdrücken sind charakteristisch für diese Behauptung: „wichtig, bedeutend, interessiert, Interesse, eine entscheidende Rolle spielen“. Danach befaßten wir uns mit Fragen, die affektive Bewertungen (a_f) enthalten. Zur Veranschaulichung geben wir einige Beispiele: „angenehm, sympathisch, zufrieden, freundlich, vertrauenerweckend, gefallen, stören, befürchten, bedauern und dergleichen“. Aus dem zuvor erwähnten ist ersichtlich, daß sich diese Fragen mit sehr spezifischen Wörtern charakterisieren lassen. Für kognitive Urteile (a_i) trifft dies jedoch nicht zu, da sich diese Urteile auf sehr unterschiedliche Themen beziehen können. Diese Urteile wurden daher erst an letzter Stelle aufgespürt.

5. In der fünften Phase wurden Verhaltenstendenzen (a_t) und Präferenzen (a_{pr}) gesucht. Diese Behauptungen sind ganz anders als kognitive Urteile und können daher relativ einfach gefunden werden. Verhaltenstendenzen sind Präferenzen, wobei man angibt, sie in der Zukunft in Taten umsetzen zu wollen. Häufig enthalten Verhaltenstendenzen sowohl charakteristische Eigenschaften von Präferenzen als auch Merkmale von zukünftigem Verhalten. Wir suchten erst nach Verhaltenstendenzen, und die Fragen die übrigblieben, waren dann automatisch Präferenzen. Typische Stichwörter für Verhaltenstendenzen sind Hilfsverben die eine zukünftige Handlung einleiten, wie etwa „sind Sie bereit ..., würden Sie..., werden Sie..., käme es in Frage, daß Sie..., und dergleichen.“ Die Verben die dann folgten gaben die spezifische Handlung an. Beispiele von Stichwörtern, die Präferenzen andeuten, sind wie folgt: „für oder gegen, dafür oder dagegen, welche Alternative bevorzugen Sie, eine Partei wählen, wählbar, zusagen, vorziehen, entscheiden, und dergleichen“.

6. Danach wurden Normen (a_n), Staatliche Maßnahmen (a_p) und Rechte (a_r) aufgespürt. Diese Behauptungen kann man unabhängig von den anderen suchen, da sie sehr spezifische Stichwörter enthalten. Normen und staatliche Maßnahmen wurden als eine Kategorie aufgefaßt, da sie dieselben Stichwörter haben wie etwa: „sollten, soll, soll werden, kann...werden, etc.“ Für Rechte wurden die folgenden Wörter verwendet: „dürfen, dürften, darf, für richtig halten, recht haben, und dergleichen“.

7. Als nächstes wurden Erwartungen zukünftiger Ereignisse (a_{ex}) gesucht mit Stichwörtern wie etwa „erwarten Sie, wird sein, und dergleichen“.

8. Danach kamen Demographische (a_d) und Verhaltensfragen (a_b) an die Reihe. Demographische Merkmale beziehen sich auf Schulbildung, Alter, Beruf, Einkommen, Haushalt, Wohnung, Mitgliedschaft von Vereinen und dergleichen. Mit diesen Stichwörtern und einigen Varianten davon konnten demographische Fragen gefunden werden. Wenn man allerdings fragt, wie verantwortlich die Stellung ist oder wieviel Einbrüche man in der Wohnung gehabt hat, dann handelt es sich nicht mehr um demographische Fragen und muß man bei der Suche nach demographischen Fragen angeben, welche Wortkombinationen auszuschließen sind.

Verhaltensfragen unterscheiden sich deutlich von demographischen Fragen. Da es sich bei den letzteren meistens um Tätigkeiten des Befragten in der Vergangenheit und der Gegenwart handelt, sind Stichwörter nötig, wie etwa „haben Sie, hatten Sie“ und Handlungsverben, wie etwa „verbringen, empfangen, betreiben, kaufen, fahren und dergleichen“.

9. In der neunten Phase wurden zusammengesetzte Behauptungen gesucht. In den Fragebögen die wir zur Verfügung hatten, kamen Sicherheit eines Urteils (a_c) und Ermittlung der Intensität einer Behauptung (a_a) vor. Für die Suche von Sicherheit eines Urteils wurden Stichwörter verwendet, wie etwa „bestimmt, durchaus, sicher, vielleicht, eventuell, etc.“ Die Anzahl der Behauptungen, die eine Ermittlung der Intensität enthielten, war sehr gering und die Fragen konnten mit den Stichwörtern „für wie gut halten Sie...“ aufgespürt werden.

10. Zuletzt wurden kognitive Urteile (a_i) gesucht. Wegen der Vielfalt der Stichwörter wurden sie erst am Ende unter Ausschließung der anderen Kategorien aufgespürt.

Nach dieser Beschreibung des halbautomatischen Suchverfahrens von Frageinhalten beschreiben wir im folgenden das Suchverfahren der Frageformen.

5.3 Die Entwicklung der computergestützten Kodierungsmethode für Frageformen

Für Frageformen entwickelten wir das folgende Suchverfahren:

1. Zuerst wurden Fragen gesucht, die sich auf einen Ort (P) beziehen mit dem Wort „wo“.

2. Danach wurden Zeitfragen (I) und ein Teil der W-Fragen aufgespürt. Für die Zeitfragen sind die Stichwörter „wann“ und „in welchem Jahr“ verwendet worden. Kausalfragen (WHY) wurden gesucht mit „warum, wodurch, worum, womit“ und mit „was waren die Ursachen“. Danach wurden die „wer“, „was“ und „wie“ Fragen mit den ent-

sprechenden Pronomen gesucht. Wenn „was“ und „wie“ in einer anderen Bedeutung vorkamen, konnten wir diese Wendungen mit einfachen Mitteln ausschließen.

3. An dritter Stelle wurden die Frageformen „welche“, „was für ein“ gesucht mit den Stichwörtern „was“ und „welche“. Da diese Kodierung nach dem zweiten Schritt erfolgt, gab es keine Probleme mit dem substantivistischen „was“, da es an dieser Stelle bereits ausgeschlossen war.

4. Die zuvor erwähnten Frageformen konnten auch eingebettet sein in Wendungen, wie etwa „Ist Ihnen aufgefallen, daß...; Erinnern Sie sich...; Kennen Sie...; Haben Sie gehört...; etc.“ Diese Fragen nannten wir Wissensfragen (K), da man ausdrücklich nach der Kenntnis des Befragten von allerlei Tatbeständen fragt. Wissensfragen waren relativ einfach aufzuspüren.

5. Nun wurden Frageformen gesucht, die Aussagen enthalten. Es geht hier um die Frageformen A und HA. A Fragen beziehen sich auf die Zustimmung mit einer Aussage oder deren Ablehnung und HA Fragen ermitteln das Ausmaß der Zustimmung oder Ablehnung. Stichwörter für A Fragen waren „Aussagen, Behauptungen, Meinungen“ in Kombination mit „zutreffen/nicht zutreffen, zustimmen/nicht zustimmen, zuneigen/nicht zuneigen“. Bei HA Fragen wurde noch nach dem Ausmaß gefragt, wie etwa „volle Übereinstimmung/teilweise Übereinstimmung, stimmt/stimmt gar nicht, trifft sehr zu/trifft etwas zu und dergleichen“.

6. Als letztes wurden die HQ und Q Frageformen aufgespürt. HQ Fragen messen, inwieweit eine Eigenschaft vorhanden ist und enthalten daher Wörter die eine Gradierung angeben, wie etwa „wie groß...wie sicher... wie zufrieden... sehr gut/eher gut/eher schlecht/sehr schlecht und dergleichen“. Nachdem die HQ Fragen gefunden waren, blieben als Restkategorie die Q Fragen übrig. Hier erwies sich unser sequentielle Verfahren wiederum als sehr nützlich, da Q Fragen sehr viele Formen annehmen können und daher schwer zu umschreiben sind.

Nach dieser Beschreibung des Klassifikationsverfahren von Frageformen, können wir im nächsten Abschnitt auf den Test des Verfahrens kurz eingehen.

5.4 Test des entwickelten Verfahrens

Das entwickelte Verfahren wurde mit der zweiten Hälfte der Fragen getestet. Das sequentielle Suchverfahren erwies sich ebenfalls als sehr brauchbar und es mußten keine Veränderungen in der Reihenfolge angebracht werden. Die Wortkriterien mußten jedoch ergänzt werden. Die meisten Ergänzungen fanden bei Frageinhalten statt, was verständlich ist, da man für die Formulierung von Begriffen über sehr viele Ausdrucks-

möglichkeiten verfügt. Aber auch bei Frageformen wurden einige Wortkriterien hinzugefügt.

Tabellen 6 und 7 vergleichen die Handkodierungen, die vom ersten Autor dieses Aufsatzes stammen, mit den Computerkodierungen anhand der Wortkriterien. Tabelle 6 bezieht sich auf die Frageformen und Tabelle 7 auf die Frageinhalte. Aus den Tabellen ersieht man, daß die Übereinstimmung zwischen den Handkodierungen und den Computerkodierungen ziemlich hoch ist. Das Computerverfahren konnte in den meisten Fällen alle Fragen entdecken. Manuell sind etwas mehr Fehler gemacht worden. Es handelt sich dabei hauptsächlich um Flüchtigkeitsfehler.

Auf Grund dieses Testergebnisses kann man schließen, daß das Verfahren brauchbar ist und daß man nach weiteren Analysen Wortkriterien erstellen können wird, die kaum noch Ergänzungen erfordern.

6. Problematische Fragen

In diesem Abschnitt wollen wir noch einige Fragen behandeln, die falsch formuliert wurden und deshalb nicht klassifiziert werden konnten. So etwa enthält Tabelle 6 eine Kategorie, die wir mehrfache Frageformen genannt haben und die noch einer kurzen Erläuterung bedarf, da sie nicht in unsere Klassifikation paßt. In den Fragebögen des Testes kamen in 14 Fällen Mehrfachfragen vor wie etwa:

„Warum haben Sie damals gerade diesen Beruf gewählt? Wer oder was hat Sie bei der Berufswahl beeinflusst?“

Mittels Handkodierung wurde festgestellt, was die Hauptfrage ist. Meistens ist es die zweite Frage, da sich die Antwortkategorien darauf beziehen und die Vorfragen als Einleitung oder rhetorische Fragen gewertet werden können. In diesem Beispiel ist es „wer oder was“, weil sich die folgenden Antwortalternativen darauf beziehen:

1. Eltern
2. Verwandte
3. Freunde
4. Lehrer
5. Berufsberater
6. Es entsprach am ehesten meinen Neigungen
7. Es ergab sich so

Wie bereits erwähnt, wurde die Hauptfrage in diesen mehrfachen Fragen durch Handkodierung festgestellt und der Kategorie „mehrfach wer oder was“ zugeordnet. Dieses

Beispiel zeigt deutlich, daß unser Verfahren halbautomatisch ist und noch manuelle Kontrollen benötigt.

Weshalb man Fragen auf diese Weise formuliert, ist nicht deutlich. Unserer Meinung nach handelt es sich um ein nachlässiges Formulieren, weil man diese Vorfragen nicht beantwortet haben will. Da sie auch Befragte verwirren können, sind diese Mehrfachfragen eher zu vermeiden.

Tabelle 7 enthält eine Kategorie mit mehrfachen Frageinhalten, die ebenfalls einer kurzen Erklärung bedarf. In fünf Fragen wurden Antwortalternativen mit verschiedenen Frageinhalten angegeben, das heißt, daß mehrere Begriffe in ein und derselben Frage verwendet wurden und sie deshalb nicht klassifiziert werden konnte, da in der Formulierung Fehler unterlaufen sind. Häufig bestehen diese Fragen auch aus mehrfachen Frageformen. Es folgen nun einige Beispiele:

„Was würde es für Sie bedeuten, wenn Sie sich morgen auf die Suche nach einem neuen Arbeitsplatz machen müßten? Welche dieser Antworten trifft am besten auf Sie zu?“

1. Das wäre eine neue Chance, ich fände schon etwas Entsprechendes
2. So gut wie jetzt ginge es mir zwar nicht, aber lange bliebe ich sicher nicht arbeitslos
3. Davor hätte ich Angst

Abgesehen davon, daß diese Frage auch zwei Frageformen enthält, nämlich „was“ und „welche“, wobei sich jedoch anhand des Stichwortes „zutreffen“ herausstellt, daß es sich um eine Zustimmung oder Ablehnung (A) handelt, enthalten die Aussagen der Antworten verschiedene Begriffe. Die erste Antwortkategorie besteht aus zwei verschiedenen Begriffen. „Das wäre eine neue Chance“ drückt eine zukünftige Erwartung (a_{∞}) aus und „ich fände schon etwas Entsprechendes“ könnte als eine Handlungstendenz (a_t) klassifiziert werden. Die zweite Antwortkategorie bezieht sich wiederum auf zwei verschiedene Begriffe. „So gut wie jetzt ginge es mir zwar nicht“ könnte als eine Evaluation (a_e) klassifiziert werden, während man „aber lange bliebe ich sicher nicht arbeitslos“ als eine evaluative Überzeugung ($a_{eb}=sD_{cy}$) betrachten könnte.

Die dritte Alternative „Davor hätte ich Angst“ ist eine affektive Bewertung $a_f=(xIf)$. Dieses Beispiel zeigt deutlich auf, daß es nicht in unsere Klassifikation paßt, da sich diese Frage auf mehrere, nicht übereinstimmende, intuitive Begriffe bezieht.

Das nächste Beispiel, das wir hier anführen ist ebenfalls problematisch:

„Wo haben Sie persönlich vielleicht ein Sparbuch, Konto, einen Kredit oder ähnliches? Ich meine, mit welchem Geldinstitut stehen Sie persönlich in irgendeiner geschäftlichen Verbindung?“

Auch hier sind wiederum zwei Frageformen verwendet worden, nämlich „wo“ und „welchem“, wobei angegeben ist, daß die zweite Frage die Hauptfrage ist. Die erste Frage besteht aus einem demographischen Merkmal $a_d=(xIc)$ „wo haben Sie ein Sparbuch...“ während die zweite Frage eine Relation enthält „mit welchem Geldinstitut stehen sie persönlich in irgendeiner geschäftlichen Verbindung“ $a_r=(sRy)$.

Ein anderes Beispiel lautet wie folgt:

„Wie ordnen Sie verschiedene Eigenschaften rein gefühlsmäßig diesen Geldinstituten zu? Nehmen wir einmal die Eigenschaft „groß und bedeutend“. Zu welchem der Geldinstitute paßt „groß und bedeutend“ am besten? Und zu welchem am zweitbesten? Und an dritter, vierter und fünfter Stelle?“

Abgesehen von den zwei Frageformen „wie“ und „welchem“, wobei die zweite Frage wiederum die Hauptfrage ist, besteht der erste Frageinhalt aus zwei Begriffen: einem Verhalten „wie ordnen Sie...zu“ $a_b=(sDy)$ und einer affektiven Bewertung mit dem Stichwort „gefühlsmäßig“. Diese Begriffskombination ist in unserer Klassifikation nicht vorhanden und kann deshalb nicht klassifiziert werden. Die Hauptfrage ist eine Präferenz: „zu welchem der Geldinstitute paßt... am besten, zweitbesten...“ $a_{pr}=(xPy,z\dots)$. Die Prädikate sind „groß und bedeutend“ und beziehen sich wiederum auf zwei verschiedene intuitive Begriffe, nämlich auf ein kognitives Urteil und auf eine Evaluation. Auch diese Begriffskombination ist in unserem Schema nicht vorhanden, da es eine falsche Formulierung betrifft, die in der Literatur als „double-barrelled“ bezeichnet wird.

Mit diesen Beispielen wollten wir darauf hinweisen, daß die Verwendung verschiedener Frageformen und Frageinhalte in ein und derselben Frage sehr verwirrend sein kann, nicht nur für die Untersucher sondern vor allem für die Befragten, was wiederum Folgen hat für die Qualität der Antworten. Wenn man unsere Klassifikation verwendet, kann man derartige Fehler schnell entdecken.

7. Zusammenfassung

Aus dieser Studie geht deutlich hervor, daß das von uns entworfene Verfahren imstande ist, Fragen nach Inhalt und Form zu klassifizieren anhand der intuitiven Begriffe, die in den Tabellen 4 und 5 beschrieben wurden. Man kann damit alle Fragen eines Fragebogens relativ problemlos klassifizieren. Zudem kann untersucht werden, ob die Fragen sich auf homogene intuitive Begriffe beziehen oder ob sie aus verschiedenartigen, intuitiven Begriffen aufgebaut sind oder ob sie Fehler enthalten. Das letztere wurde im vorigen Abschnitt erörtert.

Tabelle 6: Vergleich der Qualität der Handkodierungen mit den Computerkodierungen für Frageformen

Frageformen	Gleiche Kodierungen	Ergebnis Unterschiedliche Fehler Kodierer	Kodierungen Fehler Computer	Total
Wo (P)	5	-	-	5
Wann (I)	4	-	-	4
Warum (WHY)	5	-	-	5
Wer (WHO)	3	-	-	3
Was (WHAT)	16	1	-	17
Wie (HOW)	8	2	-	10
Welche (WHICH)	41	1	1	43
Wissen (K)	9	1	-	10
HQ	115	2	1	118
HA	18	-	-	18
A	11	-	-	11
Q	61	-	1	62
Mehrfache Frageformen	14	-	-	14
Total	310	7	3	320

Tabelle 7: Vergleich der Qualität der Handkodierungen mit den Computerkodierungen für Frageinhalte

Frageinhalte	Gleiche Kodierungen	Ergebnis Unterschiedliche Fehler Kodierer	Kodierungen Fehler Computer	Total
aeb	28	1	-	29
ae	44	8	-	52
ae	17	-	-	17
ai	9	-	-	9
aif	36	2	1	39
ai	8	-	-	8
apf	14	-	1	15
ap, an	10	-	-	10
ar	2	-	-	2
ax	11	-	-	11
ai	41	3	1	45
ab	40	2	-	42
ace	5	-	-	5
ai	2	-	-	2
a	27	2	-	29
Mehrfache Frageinhalte	5	-	-	5
Total	299	18	3	320

Da wir in Zukunft über ein effizientes und zuverlässiges Verfahren verfügen wollen, womit große Fragebestände schnell analysiert werden können, wurde ein computergestütztes Verfahren entwickelt. Die Testergebnisse dieses halbautomatischen Verfahrens waren zufriedenstellend und wir erwarten deshalb, daß mit der Zunahme unseres Fragebestandes die Wortkriterien erweitert werden können, so daß stets weniger Handkodierungen benötigt werden und das Verfahren ein hohes Maß an Zuverlässigkeit erreicht

Hiermit eröffnet sich auch die Möglichkeit, Fragebestände von großen Forschungsprojekten, die zugänglich sind (Andrews 1984, Költringer 1995, Scherpenzeel/Saris 1997), nach ihrer Qualität (Zuverlässigkeit und Gültigkeit) zu beurteilen. Man könnte dann über Kenntnis hinsichtlich der Qualität aller Fragen eines Fragebogen verfügen, bevor die Befragung stattfindet. Dies könnte eine erhebliche Qualitätsverbesserung von Fragebögen bewirken, da dieses Wissen auch bei der Formulierung von Fragen benutzt werden kann. Diese Studie sollte ein erster Ansatz dazu sein.

Derzeit ist das hier dargelegte Klassifikationsschema vor allem bei der Formulierung von Fragen nützlich. Wenn man einen Begriff gewählt hat und weiß, was man messen will, kann man mit Hilfe dieses Schemas aus den verschiedenen alternativen Formulierungen die meist geeignete Frage wählen.

Korrespondenzadresse

*Irmtraud Gallhofer, Willem E. Saris
Meander 407, 1181 WN Amstelveen, Niederlande
E-Mail: WSaris@attglobal.net*

Literatur

- Ajzen, I./Fishbein, M., 1980: Understanding attitudes and predicting social behavior. Englewoods Cliffs, N.J.: Prentice Hall.
- Ajzen, I., 1989: Attitude structure and behavior. S. 241-247 in: A.R.Pratkanis/S.J.Breckler/A.G. Greenwald (Hrsg.), Attitude structure and function, Hillsdale N.J.: Erlbaum.
- Ajzen, I., 1991: The theory of planned behavior. Organizational behavior and human decision processes 50: 179-211.
- Alwin, D. F./Krosnick, J.A., 1991: The reliability of survey attitude measurement: the influence of question and respondent attributes. Sociological Methods and Research 20: 139-181.
- Andrews, F.M., 1984: Construct validity and error components of survey measures: a structural modelling approach. Public Opinion Quarterly 48: 409-422.

- Bagozzi, R.P., 1989: An investigation of the role of affective and moral evaluations in the purposeful behavior model of attitude. *British Journal of Social Psychology* 28: 97-113.
- Blalock, H.M., Jr., 1968: The measurement problem: A gap between the languages of theory and research. S. 5-27 in: H.M. Blalock, Jr./A.B. Blalock (Hrsg.), *Methodology in the Social Sciences*. London: Sage.
- Blalock, H.M., Jr., 1990: Auxiliary measurement theories revisited. S. 33-48 in: J.J. Hox/J. De Jong (Hrsg.), *Operationalization and Research Strategy*. Amsterdam: Swets and Zeitlinger.
- Bradburn, N.M./Sudman, S., 1988: *Polls and Surveys. Understanding what they tell us*. San Francisco: Jossey-Bass.
- Brinkman, J., 1994: *Vragenlijst*. Groningen: Wolters-Noordhoff.
- Bürklin, W. P., 1980: Links und/oder Demokratisch? *Politische Vierteljahrschrift* 21: 220-247.
- Coleman, J.S., 1990: *Foundations of Social Theory*. Cambridge MA.: Belknap Press of Harvard University.
- Converse, J.M./Schuman, H., 1984: The manner of inquiry: An analysis of survey questions from across organizations and over time. S. 283-316 in: C.F. Turner/E. Martins (Hrsg.), *Surveying subjective phenomena*. Vol. 2. New York: Russel Sage Foundation.
- Cornelius, R.R., 1996: *The science of emotion. Research and tradition in the psychology of emotions*. New Jersey: Prentice Hall.
- Dahrendorf R., 1965: *Gesellschaft und Demokratie in Deutschland*. München: Deutscher Taschenbuchverlag.
- Dillon, J.T., 1984: The classification of research questions. *Review of Educational Research* 54: 327-361.
- Eagly, A.H./Chaiken, S., 1993: *The psychology of attitudes*. New York: Harcourt, Brace Jovanovich.
- Fishbein, M./Ajzen, I., 1975: *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading MA.: Addison Wesley.
- Graesser, A.C./Bommareddy, S./Swamer, S./Golding, J.M., 1996: Integrating questionnaire design with a cognitive computational model of human question answering. S. 143-174 in: N. Schwarz/S. Sudman (Hrsg.), *Answering questions: Methodology for determining cognitive and communicative survey research*. San Francisco: Jossey-Bass.
- Harary, F. 1971: *Graph theory*. London: Addison-Wesley.
- Helmers, H.M./Mokken, R.J./Plijer, R.C./Stokman, F.N. 1975: *Graven naar macht. Op zoek naar de kern van de Nederlandse economie*. Amsterdam: Van Gennep.
- Homans, G.C., 1965: *The Human Group*. 5th edition. London: Routledge and Kegan.
- Holsti, O.R., 1996: *Public opinion and American Foreign Policy*. Ann Arbor: The University of Michigan Press.
- Hox J.J. 1997: From theoretical concept to survey question. S. 47-70 in: L. Lyberg/P. Biemer/M. Collins/E. de Leeuw/C. Dippo/N. Schwarz/D. Trewin (Hrsg.), *Survey measurement and process quality*. New York: Wiley.

- Kaase M., 1971: Demokratische Einstellungen in der Bundesrepublik Deutschland. *Sozialwissenschaftliches Jahrbuch für Politik* 2: 119-326.
- Kearsley, G.P. 1976: Questions and question asking in verbal discourse: a cross disciplinary review. *Journal of Psycholinguistic Research* 5: 355-375.
- Kelley, H.H./Michela, J.L., 1980: Attribution theory and research. *Annual Review of Psychology* 31: 475-501.
- Klingemann, H. D., 1997: The left-right self-placement question in face to face and telephone surveys. *Zuma Nachrichten Spezial* 2: 113-124.
- Költringer, R., 1993: Gültigkeit von Umfragedaten. Wien: Böhlau.
- Knoke, J./Kuklinski, J.H., 1982: Network analysis. Quantitative applications in the social sciences. Vol. 28. Beverly Hills CA: Sage.
- Kretch, D./Crutchfield, S.R., 1948: Theories and problems in social psychology. New York: McGraw-Hill.
- Krosnick, J.A./Abelson, R.P., 1991: The case for measuring attitude strength in surveys. S. 177-203 in: J. Tanur (Hrsg.), Questions about questions. Inquiries into the cognitive bases of surveys. New York: Russel Sage Foundation.
- Lipset S.M. 1963: Political Man: the social basis of politics. Garden City N.Y.: Doubleday.
- Meyer, M. (Hrsg.), 1988: Questions and questioning. Berlin: de Gruyter.
- Molenaar, N.J., 1986: Formulierungseffekten in survey-interviews. Amsterdam: VU-uitgeverij.
- Northrop, F.S.C., 1947: The logic of sciences and the humanities. New York: World Publishing Company.
- Oppenheim, A.N., 1966: Questionnaire design and attitude measurement. London: Heinemann.
- Oskamp, S., 1991: Attitudes and opinions. 2nd edition. Englewoods Cliffs N.J.: Prentice Hall.
- Parsons, T., 1951: The social system. Glencoe Ill.: Free Press.
- Rabinowitz, G./Macdonald, S.E./Lishuag, O., 1991: New Players in an Old Game: Party Strategy in Multiparty Systems. *Comparative Politics* 24: 147-185.
- Rokeach, M., 1973: The nature of human values. New York: Free Press.
- Saris, W.E., 1998: The split ballot MTMM experiment: An alternative way to evaluate the quality of questions (Manuskript zur Publikation angeboten).
- Saris, W.E./Gallhofer, I.N., 1997: Formulation and Classification of questions (Manuskript zur Publikation angeboten).
- Saris, W.E./Gallhofer, I.N., 1998: Classificatie van survey-vragen. *Tijdschrift voor Communicatiewetenschappen* 26: 98-121.
- Scherpenzeel, A.C./Saris, W.E., 1997: The validity and reliability of survey questions: A meta-analysis of MTMM studies. *Sociological methods and Research*. 25: 341-383.
- Schuman H./Presser, S., 1981: Questions and answers in attitude survey: experiments on question form, wording and context. New York: Academic Press.

- Smith, T.W., 1987: The art of asking questions 1936-1985. *Public Opinion Quarterly* 51: 95-108.
- Sniderman, P.M./Brody, R.A./Tetlock, P.E., 1991: *Reasoning and Choice: Explorations in Political Psychology*. Cambridge MA: Cambridge University Press.
- Sorokin, P., 1928: *Contemporary sociological theories*. New York: Harper.
- Sudman, S./Bradburn, N.M./Schwarz, N., 1982: *Asking questions. A practical guide to questionnaire design*. San Francisco: Jossey-Bass.
- Tesser, A./Martin, L., 1996: The Psychology of Evaluation. S. 400-432 in: E.T. Higgins/A.W. Kruglanski (Hrsg.), *Social Psychology. Handbook of basic principles*. New York: Guilford Press.
- Torgerson, W. S., 1958: *Theory and methods of scaling*. New York: Wiley.
- Van der Pligt, J./De Vries, N.K., 1995. *Opinies en attitudes. Meting, modellen en theorie*. Amsterdam: Boom.
- Von Winterfeldt, D./Edwards, W., 1986: *Decision analysis and behavioral research*. Cambridge MA: Cambridge University Press.
- Zanna, M.P./Rempel, J.K., 1988: Attitudes: a new look at an old concept. S. 114-125 in: D. Bar-Tal/A. Kruglanski (Hrsg.), *The social psychology of knowledge*. Cambridge MA: Cambridge University Press.