

Forschungstechniken als Teil der Soziologie heute

Scheuch, Erwin K.

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Scheuch, E. K. (1976). Forschungstechniken als Teil der Soziologie heute. In M. R. Lepsius (Hrsg.), *Zwischenbilanz der Soziologie: Verhandlungen des 17. Deutschen Soziologentags* (S. 83-127). Stuttgart: Ferdinand Enke. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-160661>

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.

III. Methodologie und Forschungstechniken

Forschungstechniken als Teil der Soziologie heute

Erwin K. Scheuch

1.

Im Erscheinungsbild der Öffentlichkeit, nicht zuletzt auch dem an manchen Universitäten, scheinen Methodenfragen die Soziologie als eine Disziplin zu spalten. Ein Vergleich der tatsächlichen Lehrangebote an Universitäten nicht nur der Bundesrepublik, sondern auch anderer westeuropäischer Länder, ergibt ein nahezu gegenteiliges Bild. Wie verschieden auch im weiteren Verlauf des Studiums die Lehrpläne ausschauen, wie unterschiedlich auch die inhaltlichen Positionen des Personals an verschiedenen Hochschulen sein mögen: Eine Grundausbildung in zumindest Techniken der Sozialforschung gehört zu den verbliebenen Gemeinsamkeiten einer in vielen anderen Fragen tatsächlich uneinigen Disziplin.

Wahrscheinlich ist dies mehr als eine verbliebene Gemeinsamkeit, wahrscheinlich wird es zunehmend überhaupt keine Streitfrage mehr sein, empirisch ausgebildet zu sein und zu arbeiten. Die noch weiter geführte methodologische Grundsatzdiskussion wird immer weniger Bezug zur Praxis der Soziologie als Forschung haben (1). Ohnehin war dieser Bezug schwach ausgebildet. Das gilt auch für die Vertreter methodologischer Positionen, die in der Bundesrepublik mit der irreführenden Bezeichnung "Positivisten" belegt wurden (2). Welche Bedeutung hat denn je das Falsifikationsprinzip für die tatsächliche Forschung gehabt? Gibt es eine irrelevantere Veröffentlichung für die Entwicklung der Forschung als das Buch mit dem Titel "Positivismusstreit" (3)?

Ungeachtet der Absichten, die sich mit empirischer Forschung verbinden, wird die Technologie der Sozialforschung – neben einer gemeinsamen Begrifflichkeit zur Kennzeichnung unmittelbar beobachtbaren Verhaltens – zu einem der die Soziologie als Profession kennzeichnenden Elemente (4). Ein Vergleich mit der politischen Wissenschaft zeigt, welche Bedeutung einer solchen eher handwerklichen Dimension für die Soziologie als Disziplin zukommt.

Voraussetzung für diese Bedeutung der Forschungstechniken ist die handwerkliche Standardisierung. Hierfür eignen sich nicht alle Forschungstechniken in gleicher Weise – zumindest nicht im gegenwärtigen Stand ihrer Ausbildung. Erst mit der formelhaften Standardisierung einer Technik kann diese in einen Standardkatalog von Lehr- und Prüfungsstoff auf den elementaren Stufen einer soziologischen Ausbildung aufgenommen werden (5).

Wenn den Forschungstechniken der Soziologie für das Selbstverständnis und die Präsentation der Soziologie in der akademischen Öffentlichkeit eine so zentrale Bedeutung zukommt, dann war in einem Augenblick des Übergangs der Soziologie zu einem Massenlehrfach eine Bestandsaufnahme angezeigt – eine Bestandsaufnahme der Soziologie in ihrer Erscheinungsform als Kunstlehre. Hierbei erwies es sich schon bald, daß es kein einheitliches Erklärungsobjekt für eine solche Bestandsaufnahme gibt: Die Kodifizierung der Forschungstechniken als Teil einer

Standardausbildung hat einen anderen Charakter als die praktische Forschung, und diese wiederum differiert in ihren Vorgehensweisen von der methodologischen Literatur.

Zwischen der Forschungstechnologie in diesen drei Erscheinungsformen und den inhaltlichen Fragestellungen besteht nur ein loser Zusammenhang. Das war zu erwarten für Forschungstechniken als Teil eines standardisierten und abprüfbaren Lehrstoffes; dieser Zustand ist dem in anderen etablierten Disziplinen nicht unähnlich. Da der Zusammenhang zwischen Lehre und Soziologie als wissenschaftliches Fach hier nicht Thema ist, soll die Lehre der Forschungstechniken – wie verselbständigt diese auch gegenüber der praktischen Forschung sein mag – hier nicht weiter besprochen werden. So verbleibt als Gegenstand der Betrachtung die Art der Beziehungen zwischen der praktischen Forschung und der Methodologie als einer verselbständigte Spezialität. Drei Themen sollen in diesem Übersichtsreferat nur kurz erwähnt werden: (1) Ein neuer Ansatz in der Inhaltsanalyse, jetzt verstanden als eine Form der Verbindung von Aussagen der Alltagssprache mit der begrifflichen Umsetzung in eine Fachsprache und unter Verwendung moderner EDV Techniken (*Stone* und *Mochmann*); (2) eine Neubelebung des Interesses an nicht-reaktiven Forschungstechniken und an dem Indikatorwert von Daten (*Lück*); (3) als Kontrast zur Darstellung der heutigen Methodenlehre die Kennzeichnung der tatsächlichen Forschung in der Bundesrepublik (*Herz* und *Stegemann*). Insgesamt sollte dann aus dem Ensemble der Beiträge deutlich werden, daß sowohl inhaltliche Interessen der Forschung stimulierend auf die Entwicklung der Forschungstechnologie wirken, wie umgekehrt technologische Entwicklungen neue inhaltliche Fragestellungen provozieren. Diese Wechselbeziehung wird in der wissenschaftstheoretischen Grundsatzdiskussion wie auch in der scholastisch verfestigten wissenschaftsphilosophischen Diskussion über Forschung unterbewertet (6).

Es ist nicht Absicht dieses Trendberichts, Begründung für irgendwelche Apelle zu sein. Es sollen Sachverhalte deutlich gemacht werden – insbesondere der Sachverhalt einer wesentlich loseren Beziehung zwischen Methodologie und praktischer Forschung als oft unterstellt wird, es soll aufmerksam gemacht werden auf einige besonders wichtige methodologische Entwicklungen und auf einige Veränderungen in der Problemstellung von soziologischer Forschung. Letzteres führt hin zu der abschließenden Frage, die weitgehend Frage bleiben muß: Der Beziehungen zwischen Soziologie als einer inhaltlichen Disziplin und soziologischer Forschung. Daß hierbei auf Soziologie als einer Erfahrungswissenschaft unter anderen Erfahrungswissenschaften abgestellt wird, ist nach dieser Charakterisierung des Trendberichts selbstverständlich.

2.

Die Kodifizierung der Forschungstechniken und die Routinisierung der Sozialforschung kamen bekanntlich nach Ende des 2. Weltkriegs als Übernahme amerikanischer Entwicklungen nach Europa. Hierdurch entstand das Bild, die Empirie sei in unserem Fach eine "amerikanische Soziologie" (7). In Korrektur dieses auch heute noch verbreiteten Klischees konnte darauf verwiesen werden, daß die Mehrzahl der wichtigen Forschungstechniken zumindest als Prototyp zunächst in

Europa entwickelt wurden: Der standardisierte Fragebogen (8), Beobachtungsverfahren (9), Stichprobenverfahren (10), Inhaltsanalysen (11), Psychotests (12), Soziometrie (13), Einzelfallstudie (14). Die wirklichen Eigenentwicklungen in den USA bezogen sich in erster Linie auf Analysetechniken (15).

Dennoch ist die Zurückweisung irriger Vorstellungen über Prioritäten in einem anderen, hier relevanten Zusammenhang ziemlich unerheblich. Die Sozialforscher in Europa blieben tatsächlich in der Zeit seit dem 2. Weltkrieg gegenüber der Entwicklung in der Methodologie weitgehend rezeptiv. Eigenständige Beiträge beschränkten sich auf einige, allerdings wichtige, Sondergebiete wie die Mehrebenenanalyse oder den interkulturellen Vergleich. Bezeichnenderweise erschienen diese Beiträge überwiegend in englischer Sprache und wurden durchweg in den USA stärker beachtet als in den jeweiligen europäischen Ländern selbst (16). Methodologie als Spezialität und eine entsprechende professionelle Öffentlichkeit wurde eben in den USA ausgebildet – dort schon zu Beginn der Sozialforschung (17). Von daher beziehen wir in Europa auch heute noch diejenigen Weiterentwicklungen in der Methodologie, die über das Stadium prototypischer Neuerungen hinausgehen.

Diese Charakterisierung ist nicht zu verstehen als Klage über eine ungenügende Verbreitung von Kenntnissen in der Methodologie und den Techniken in der Sozialforschung. Sozialforschung als Teil der Lehre und als praktisches Vorgehen existiert in den USA auf sehr unterschiedlichen Niveaus. Wird nicht lediglich das als Maßstab genommen, was bei guten amerikanischen Zeitschriften den Filter kritischer Beurteilungen der beratenden Redakteure passiert, sondern werden in die Betrachtung die insgesamt eingereichten Manuskripte mit einbezogen, so erweist sich ein erheblicher Teil der Forschungen nicht nur als banal und unzureichend an der theoretischen Diskussion orientiert, sondern auch als handwerklich ungenügend (18). Andererseits entspricht – nach den für die Lehre verwandten Unterlagen zu urteilen – die Ausbildung in Methodik der empirischen Sozialforschung an den für unser Fach wichtigeren Hochschulen der Bundesrepublik in etwa den Standards an vergleichbaren Universitäten der Vereinigten Staaten, oder ist doch zumindest an diesen Standards orientiert (19). Der kritische Unterschied besteht einmal in der kleinen Zahl eines methodologisch und technisch wirklich hoch qualifizierten Personals in Westeuropa und in dem Verzicht der meisten Hochschullehrer, nach den Anfangsphasen ihrer Karriere weiter selbst empirische Sozialforschung zu betreiben (20). Diese beiden kritischen Umstände können in diesem Zusammenhang nicht kausal abgeleitet werden; sie seien lediglich konstatiert als mit-ursächlich dafür, daß die Methodologie in der Bundesrepublik – und übrigens auch in den anderen Ländern Europas – keine Spezialität als eigenes Sachgebiet wird.

Dieser Aussage scheint zunächst zu widersprechen, daß gerade in den letzten Jahren in der Bundesrepublik eine beachtliche Zahl größerer Darstellungen insbesondere zu forschungstechnischen Frage erschienen. Dazu gehören durchaus eigenständige Untersuchungen zu den Fehlerquellen des Interviews (21), zu Beobachtungstechniken (22) und zu Meßverfahren (23); in einigen Fällen wurden Themen bearbeitet, für die gegenwärtig in anderen Sprachen kein Äquivalent vorliegt (24). Die Entwicklung in Analyse-Techniken insbesondere in den USA wurde in teilweise eigenständiger Weise aufgearbeitet, wengleich diese Mono-

graphien eher rezeptiv blieben (25). Eine Monographien-Reihe, die seit Anfang der siebziger Jahre erscheint, soll die Lehre der Forschungstechniken für die Bundesrepublik auf einem Niveau standardisieren, das dem eines Bachelor of Arts einer amerikanischen Universität mit einer überdurchschnittlich guten Ausbildung in Sozialforschung entspricht (26). Vor allem aber ist das "Handbuch der empirischen Sozialforschung" zu erwähnen, das soeben in dritter und stark überarbeiteter Form erscheint (27). Dieses "Handbuch" hat seit seiner ersten Auflage 1962 Forschungstechniken und Methodologie in einer Breite und dennoch miteinander verbunden vorgestellt, dem in anderen Sprachen höchstens das "Handbook of Social Psychology" entspricht (28). Hinzu kommen eine Reihe durchaus brauchbarer Einführungen in die Sozialforschung (29). Heute ist es möglich, sich – mit einigen nicht zentralen Ausnahmen – in deutscher Sprache über denjenigen Stand der Sozialforschung, über den Konsens unter Methodologen besteht, unmittelbar zu unterrichten.

Drei weitere Aspekte müssen für eine angemessene Kennzeichnung der Situation in der Bundesrepublik jedoch ebenfalls berücksichtigt werden: 1. Die Verfasser der Einzeldarstellungen stehen durchweg noch am Beginn ihrer Karriere, können deshalb nur wenig Forschungserfahrungen haben, sodaß die Darstellungen den Charakter von Buchwissen-Empirie meist nicht vermeiden können; 2. die weitaus meisten Publikationen über Methodologie i.e.S. haben fast keinen Bezug zur Sozialforschung, sondern sind "philosophy of science" oder deren Bekämpfung (30); 3. die Fachzeitschriften als das wichtigste Forum einer Fachwissenschaft sind für die Entwicklung bei den Methodenfragen von nur nebensächlicher Bedeutung (31). Letzteres ist sicherlich keine Frage der Redaktionspolitik, sondern eine Folge des geringen Angebots an entsprechenden Manuskripten.

Wer die Medien und das Publikum für Methodenfragen der Sozialforschung empirisch abgrenzen will, dem sei eine Auswertung der Fußnoten bzw. Bibliographien der im vorigen Abschnitt angesprochenen Veröffentlichungen empfohlen. Die große Mehrzahl aller Autoren, auf die sich die Veröffentlichungen in der Bundesrepublik beziehen, sind hiernach Amerikaner und die wichtigsten Primärarbeiten erscheinen in einigen wenigen amerikanischen Zeitschriften. Eine Kennzeichnung der Entwicklungen bei Methodenfragen der Sozialforschung, die sich auf Primärquellen stützen sollte, muß sich auf die Publikationen konzentrieren, in denen die wichtigsten Primärarbeiten erscheinen – d.h. solche, die wiederholt als Grundlage für weitere Arbeiten dienen. Aus dieser Abgrenzung folgte für diesen Trendbericht, daß die folgenden Quellen systematisch auszuwerten waren (32):

- American Journal of Sociology – ab 1968/9 (Band 74)
- American Sociological Review – ab 1969 (Band 34)
- Journal of Mathematical Sociology – ab 1971 (Band 1)
- Kölner Zeitschrift für Soziologie und Sozialpsychologie (33) – ab 1969 (Band 21)
- Sociological Methods and Research – ab 1972 (Band 1)
- Sociological Methodology – Sammelbände ab 1969
- Sociometry – ab 1970 (Band 33)
- Public Opinion Quarterly – ab 1968 (Band 32)

Näher angeschaut wurden lediglich Artikel mit methodischem Schwerpunkt. Die Kurzbemerkungen ("Comments" und "Rejoinders") blieben für diese Auswertung durchweg unberücksichtigt, ebenso wie methodische Ausführungen im Rahmen eines primär anderen Themen gewidmeten Artikels. Damit verblieben aus den sieben

zunächst erwähnten, als international etabliert geltenden Medien für eine Untersuchung 59 Artikel. Das Public Opinion Quarterly als achte der Zeitschriften veröffentlichte demgegenüber 115 als relevant ausgewählte Aufsätze. Diese beiden Kollektive im Umfang von 59 plus 115 sind die Gesamtheit des Materials, das hier als bestimmend für die internationale methodologische Diskussion untersucht wird.

Wahrscheinlich hätte die Grenze des zu berücksichtigenden Kollektivs auch weiter gezogen werden können. Zeitschriften wie "Sociology and Social Research" oder "Sociological Review", "Soziale Welt" oder das "Journal of Marketing" bringen beachtliche Artikel. Ihr jeweiliges Publikum ist aber spezieller als selbst das des Public Opinion Quarterly. Entscheidender war jedoch die Überlegung, daß bei Zweit- oder Drittveröffentlichungen über die gleiche Untersuchung, mindestens eines der als zentral für die methodisch interessierte Öffentlichkeit ausgewählten Medien bevorzugt zu werden pflegt.

Bewußt wurde auf eine Auswertung von "Readern" verzichtet. Dieser Typ von Veröffentlichung wäre relevant gewesen, wäre versucht worden, den bereits kodifizierten Teil der methodologischen und forschungstechnischen Erörterungen zu kennzeichnen (34). Hier sollte jedoch die aktuelle Entwicklung als professionelle Kommunikation gekennzeichnet werden.

Der Erhebungszeitraum ergab sich aus einer Zielsetzung dieses Trendberichts: Untersucht werden sollte u.a. der Bezug zwischen professioneller Diskussion und Erörterung der Soziologie in einer allgemeinen Öffentlichkeit. Ab 1967 hatte es den Anschein, als ob die Profession kein entscheidender Kommunikationszusammenhang mehr wäre; Soziologie schien durch eine Erörterung in einer allgemeinen, intellektuell interessierten Öffentlichkeit in Selbstverständnis und Themenwahl gesteuert zu werden. Dies hätte sich in den verschiedenen Zeitschriften unterschiedlich schnell ausdrücken müssen, da in ihnen die Zeit zwischen Beendigung eines Aufsatzes und Veröffentlichung bis zu zwei Jahren differiert.

Wenn damit nach Zeitraum und Umfang das Kollektiv so abgegrenzt wurde, daß es als Forum professioneller Diskussion über Methodenfragen behandelt werden darf, wäre eine Auswertung primär nach Häufigkeiten (z.B. Prozentsatz der Erwähnung von Faktorenanalysen in Relation zur Pfadanalyse) nicht angemessen. Damit würde die bloße Häufigkeit der Themenwahl vielleicht abgebildet, aber auch dies auf eine nur schwierig zu interpretierende Weise. Die Artikel sind nämlich als Zählheiten unterschiedlich unabhängig voneinander: Nicht selten wird von der gleichen Forschergruppe ein Thema in Raten behandelt, in anderen Fällen bilden die Aufsätze eine Argumentationskette von aufeinander reagierenden Autoren. Es wäre sehr lehrreich zu verfolgen, wie ein Aufsatz auf andere, später veröffentlichte einwirkt, oder ob eine zunächst isolierte Veröffentlichung auch über den weiteren Beobachtungszeitraum ohne Reaktion bleibt (35). Mit diesen letzteren Überlegungen wird auf einen zentralen Gesichtspunkt bei einer Auswertung von Literatur verwiesen, nämlich das unterschiedliche Gewicht von Aufsätzen; dies ist jedoch bei fehlendem zeitlichen Abstand nur ungenügend erfassbar. Aus diesen Erwägungen heraus wurde im Gegensatz zu manchen amerikanischen Veröffentlichungen die Aufgabe dieses Trendberichts darin gesehen, statt Aufzählungen zu bringen, Schwerpunkte und Tendenzen aufzuzeigen (36).

3.

Zu Beginn unserer Beobachtungsperiode wurde in einer repräsentativen Veröffentlichung in den USA eine "Revolution in soziologischen Methoden" als unmittelbar bevorstehend vorausgesagt (37). Die erste Revolution – so der Autor – ereignete sich in den dreißiger Jahren mit der Entwicklung der Empirie, so wie sie uns heute vertraut ist. Literarisch habe sich diese "Revolution" in einer Verdrängung des Essays durch den technischen Forschungsbericht ausgewirkt, der seither die amerikanischen Zeitschriften kennzeichne. "For the greater part, this is where matters stand today. The sociologist is a gatherer of facts. His basic working tools are the interviewing techniques of survey research and statistical analysis. His greatest methodological concerns are with the validity and reliability of his observations and with related problems of measurement" (38). Die bevorstehende "zweite Revolution" sei gekennzeichnet einmal durch die Mathematisierung der Forschung und zum anderen durch die wirksame Nutzung von Großcomputern.

Während in einer ungefähr gleichzeitigen deutschen Publikation den Großcomputern eine erstrangige Bedeutung beigemessen wurde, sieht der amerikanische Autor in der Mathematisierung die tiefgreifendere Veränderung (39). Zu recht verweist der amerikanische Autor auf die Rückwirkungen einer Mathematisierung auf Themenwahl und Problemformulierung. Die uns bisher bekannten Formen einer Mathematisierung eignen sich nicht sonderlich zur Analyse komplexerer Phänomene, und entsprechend wird von diesen mathematisch orientierten Sozialforschern die Untersuchung von elementar-vereinfachten Gegenständen bevorzugt. Tauschbeziehungen, Cliquenverhältnisse, Machtverhältnisse in Kleingruppen – es sind bevorzugt Gegenstände, die sich als Interaktionen erfassen lassen, und die selbst dann noch definitorisch radikal vereinfacht werden (40). Es werden gewiß auch Makrophänomene untersucht, wie Wanderungsbewegungen oder vertikale Mobilität, aber im Schwerpunkt handelt es sich doch um Themenstellungen, die auch experimentellen Versuchsanordnungen zugänglich sind (41). Von den Befürwortern dieser mathematischen Verfahren wird immer wieder behauptet, man werde später von elementaren zu komplexen Phänomenen fortschreiten, aber bisher hat sich diese Richtung als wohlentwickelte Spezialität "mathematische Soziologie" in das Spektrum soziologischer Spezialitäten eingeordnet, statt generell stilbildend zu wirken.

Die Auswertung der Zeitschriften und anderen Quellen ergibt – mit der wichtigen Ausnahme des "Public Opinion Quarterly", – daß *Robert McGinnis* den Schwerpunkt der Aufsatzproduktion zur Methodologie der Sozialforschung richtig voraussah. Die Mathematisierung ist weiter fortgeschritten – allerdings keineswegs so stürmisch, daß die Bezeichnung "Revolution" auch nur annähernd gerechtfertigt wäre, und zudem im Schwerpunkt als mathematische Statistik und nicht in Richtung auf "mathematische Soziologie" i.S.d. Konstruktion von Modellen. Gewiß ist damit nun auch in der Soziologie ein Zustand eingetreten, der in der Nationalökonomie seit längerem besteht: Die Forscher eines Faches können nicht mehr alle Aufsätze in einer allgemeinen Zeitschrift verstehen; dennoch handelt es sich auch bei den "unverständlicheren" Aufsätzen um eine Fortführung bereits bekannter Themen. Vier Problemstellungen sind für die amerikanischen Aufsätze hierbei besonders wichtig: Die Fortführung der Erörterung über Maße zur Bestimmung von Korrelationen; der sich teilweise als Alternative hierzu verstehende

Versuch der Aufdeckung von Kausalität durch Pfadanalyse und ähnliche Ansätze; Probleme des Abbildens einer latenten Eigenschaft durch Messen, insbesondere mehrdimensionale Skalierungen; und die Diskussion um die Angemessenheit parametrischer versus nicht-parametrischer Maßzahlen.

Bereits mit diesen allgemeinen Beobachtungen lassen sich einige Schlußfolgerungen über die Entwicklungsrichtung der methodologischen Literatur begründen. Es ist offensichtlich, daß Niemand mehr in dem umfassenden Sinne allgemeiner Methodologe sein kann, wie dies noch in den fünfziger Jahren möglich war, nämlich als Fachmann, der über die jeweiligen neuesten Entwicklungen auf allen wichtigen Verfahrensgebieten ein durch eigene Sachkenntnis begründetes Urteil hatte. Insbesondere die Analysetechniken haben sich gegenüber den anderen Gebieten der Methodologie zunehmend verselbständigt. Noch 1967 konnte geschrieben werden: "In der Sozialforschung hat seit Beginn der 60er Jahre eine Akzentverlagerung stattgefunden, die sich in der methodologischen Literatur bisher nur unvollkommen ausdrückte. . ." (42). Gemeint war die Akzentverlagerung der Aufmerksamkeit von Forschern zur Analyse hin. Diese Akzentverlagerung ist inzwischen voll in der Literatur durchgeschlagen, ja diese Literatur hat in ihren Verästelungen und Spezialisierungen die Forschung "überholt" – falls man in einer bloßen Differenzierung von Verfahren einen Fortschritt der Forschung sieht.

Am ehesten entsprechen die Aufsätze zur Pfadanalyse und verwandten Kausalanalysen den Vorstellungen über die Mathematisierung. Die Vertrautheit mit dem Verfahren wird vorausgesetzt und die Autoren stellen Techniken zur Lösung spezifischer Probleme vor. So ist es öfters nützlich, statt der Vielzahl von Koeffizienten, die Kausalität in einer Pfadanalyse ausdrücken, ein zusammenfassendes Maß zu haben. Ein solches Maß wird von *David R. Heise* mit dem "sheaf coefficient" vorgestellt, bei dem zugunsten des Einflusses der für den Ausdruck der Kausalität interessierenden Variablen die anderen Variablen unterdrückt werden (43). Es wird dabei allerdings nicht die an sich naheliegende Frage untersucht, ob nicht bei einem Bedürfnis nach einer einzelnen zusammenfassenden Maßzahl, die zudem durch Unterdrückung von Variablen gewonnen wird, auf ein anderes Verfahren zurückgegriffen werden sollte, da ja auf den besonderen Vorzug der Pfadanalyse – die differenzierte Darstellung der relativen Bedeutung direkter und indirekter Faktoren – verzichtet wird. Der Berechnung des Beitrags von Faktoren, die über andere Faktoren auf eine Erscheinung einwirken, ist das Thema des Aufsatzes von *John M. Finney*, in dem er nachweist, daß solche indirekten Effekte bisher nur unsauber berechnet wurden (44). Die Nutzung von Faktorenanalysen (allerdings unter spezifischen Bedingungen) bei der Konstruktion von Modellen für Pfadanalysen, ist das Thema eines Aufsatzes dreier Autoren (45). Abgesehen von seiner rechentechnischen Detailliertheit ist an diesem Aufsatz auch charakteristisch, daß er ein recht technisches Einzelthema des Hauptautors *Karl W. Jöreskog* über Jahre hinweg weiterführt. In den beiden Aufsätzen von *Erich Weede* in der "Kölner Zeitschrift" kann demgegenüber diese Vertrautheit mit Pfadanalyse nicht unterstellt werden, sodaß dieser Artikel den Charakter eines Lehrbuchaufsatzes erhält (46).

Besonders vielfältig verästelt, spezialisiert bis zu teilweise Unverständlichkeit auch für statistisch Ausgebildete, und unübersichtlich waren die Beiträge, in denen neue Maßzahlen empfohlen werden. Besonders häufig werden Korrelationsverfah-

ren vorgestellt, die besser als die allgemein verbreiteten Maßzahlen dem besonderen Charakter der Daten gerecht werden sollen. *Roland H. Hawkes* schlägt wieder einmal Vorgehensweisen für Partialkorrelation vor, die dem üblichen Meßniveau in der Sozialforschung – ordinalem Messen – besser als die unter Soziologen beliebten Koeffizienten entsprechen sollen (47). Ordinales Meßniveau wird auch in anderen Beiträgen zur Messung von Covariation vorausgesetzt, wobei durch einfache Ausdrucksform der Variablen (Dichotomien oder Trichotomien) Maßzahlen konstruierbar werden, welche die Determinationskraft von unabhängigen Variablen so wiedergeben sollen, wie dies für Intervallmessungen etwa durch " r^2 " geschieht (48). Wie aus Daten eines niedrigen Meßniveaus dennoch inhaltsreichere Aussagen als mit bisher bekannten Verfahren erreicht werden können, ist das Thema von *Tai S. Kang* (49). Er erreicht dies, indem er zur Stärke des Zusammenhanges zwischen zwei Variablen – über die naturgemäß auf ordinalem Meßniveau nicht allzu präzises ausgesagt werden kann – noch die Form des Zusammenhangs ausdrücken will. Unter Rückgriff auf ein älteres, aber wenig beachtetes Verfahren der "multiple component analysis" – also des ex-post Auseinandersortierens von Einflußbündeln – wird ein Verfahren vorgestellt, das auf der Grundlage von Vektoren bzw. Matrizenrechnung Analoges bewirken soll wie eine Pfadanalyse (50).

Besonderes Interesse fanden im Zusammenhang mit Verfahren zum Aufweis von Abhängigkeiten die Anwendungsmöglichkeiten von Markov-Ketten. Während die Mehrzahl der Aufsätze zu Covariationsproblemen sonst von einem bestimmten Sachgebiet abheben, überwiegt hier Darstellung an den Problemen eines Anwendungsbereiches, der sozialen Mobilität im Generationsverlauf (51). Wahrscheinlich ist hier der Umstand besonders attraktiv, daß die Gerichtetheit des Prozesses, der durch Markov-Ketten modelliert werden soll, eindeutig ist, wie andererseits eine Anwendung dieses Modells ohne Modifikation sicherlich sachlich nicht angemessen sein dürfte. Wie das wegen seiner bekannten mathematischen Eigenschaften für nicht-Mathematiker interessante Kalkül doch auf inhaltlich vertretbare Weise angewandt werden kann, ist das durchgehende Problem.

"Klassische" Themen der statistischen Methodenlehre für Sozialwissenschaftler kommen überraschend selten vor. Dazu gehört die Messung von Korrelation beim Test-Retest-Verfahren – ein bei der Testkonstruktion entscheidender Schritt – für die während des Beobachtungszeitraums ein neues Verfahren vorgeschlagen wird, durch das Verlässlichkeit und Stabilität der Ergebnisse zwischen verschiedenen Messungen im Zeitablauf rechnerisch getrennt ausgedrückt werden können (52). Ein "klassisches" Thema war auch der statistische Ausdruck von Verteilungsformen von Daten. Hier soll durch ein weiterentwickeltes Verfahren mittels Standardisierung ein Vergleich zwischen verschiedenen Verteilungen erreicht werden (53). Da wir in der Sozialforschung für diesen Zweck bereits über Maßzahlen verfügen, wäre es besonders nützlich gewesen zu erfahren, warum wir nun diese Maßzahl benötigen. Es ist diese Vervielfältigung von Maßzahlen für teilweise höchst ungewöhnliche Sonderfälle, die kürzlich den Methodologen *Otis Dudley Duncan* zu der Forderung nach einem "Moratorium für Methodologie" bewog (54). *Duncan* beklagte, daß die Mehrzahl der neuen Maßzahlen lediglich "advertise some new procedure for massaging numbers"; sie seien aber nicht aus dem Bedürfnis heraus entwickelt worden, ein Sachproblem zu lösen. So habe man heute inkompetente praktische Forscher, die blindlings der vorgeblichen Autorität eines

“Methodologen” folgten und andererseits Methodologen, deren Rat zu nichts taue, weil sie sich nicht mit Forschung als einer Klärung von Sachverhalten beschäftigten (55).

Von diesem Vorwurf nimmt *Duncan* einen Methodologen ausdrücklich aus, der in den Zeitschriften und Lehrbüchern außerhalb der USA noch wenig beachtet wird, obgleich er die Analyse in der Sozialforschung unter Verwendung moderner Computern nachhaltig beeinflusste – *Leo A. Goodman* von der Universität Chicago. *Goodman* wurde besonders bekannt durch den Entwurf verschiedener Modelle zur Analyse von Umfragedaten. Während des Beobachtungszeitraums veröffentlichte er in diesem Zusammenhang ein Modell, durch das eine Kausalerklärung für Umfragedaten analog Pfadanalysen oder verwandter Verfahren möglich werde – also ein Nachweis der relativen Stärke und Art der Interaktion zwischen mehreren Variablen (56). Es ist kennzeichnend für *Goodman*, daß er dieses Modell an der Sekundäranalyse eines Materials – in diesem Falle über die These, daß Alltagskontakte Distanz zwischen Schwarzen und Weißen verringerten – entwickelt und zugleich mit der Demonstration des Verfahrens zu zusätzlichen inhaltlichen Aussagen kommt.

Dies trifft auch für den etwas früheren Aufsatz über Analyse-Verfahren für Tabellen zu, deren Nützlichkeit er an der Analyse von Daten über Generationsmobilität aufzeigt (57). In diesem Falle kann *Goodman* als Korrektur zu bisher vorherrschenden Ansichten nachweisen, daß in England und Dänemark heute die soziale Auf- und Abwärtsmobilität für Kinder aus mittelständischen Soziallagen am höchsten ist. *Leo Goodman* hat entscheidende Beiträge zur Analyse von Panel-Daten geliefert, was den Vertretern dieser Art von Versuchsanordnung, den Sozialwissenschaftlern um *Paul F. Lazarsfeld*, nicht gelang. Während des Beobachtungszeitraums legt *Goodman* hierzu einen weiteren Beitrag vor, in dem an den Daten über Wahlabsicht aus der berühmten Wahlstudie von *Paul Lazarsfeld* “The People’s Choice“ eine Technik der Kausalanalyse erläutert (58). Bedeutend sind auch seine Beiträge zu den Skalierungsverfahren von *Guttman* und der “Analyse latenter Strukturen“ von *Lazarsfeld*, die bisher nur als mathematisches Kalkül aber ohne die notwendigen statistischen Verfahren vorgestellt wurden; diese statistische Begründung gelingt dem ausgebildeten Mathematiker *Goodman* auch in einer Publikation über indirektes Messen während des Beobachtungszeitraums (59). In die praktische Forschung sind aber heute schon vor allem *Goodman’s* Beiträge zum Ausdruck von Covariation eingegangen. In einem Beitrag zur Messung von Kausalität bei Umfragedaten befaßt sich *Goodman* während dieses Zeitraums mit Regressionsanalysen für Tabellen mit dichotomisierten Variablen, wobei wiederum die interaktive Wirkungen mehrerer Variablen nachweisbar werden (60). In allen diesen Fällen resultiert die Sekundäranalyse mit dem jeweils neu vorgestellten Verfahren, für die *Goodman* Untersuchungen bevorzugt, welche in die Standardliteratur der Soziologie eingegangen sind, in inhaltlich neuen Aussagen – teilweise zusätzliche Aussagen, teilweise auch Korrekturen für bisher als belegt geglaubte Sätze.

Selbstverständlich werden in den beobachteten Quellen auch weiterhin die früher vorherrschenden Themen der methodologischen Literatur behandelt. Eines der wichtigsten Themen war in der ersten Hälfte der sechziger Jahre die Skalierung, und dabei insbesondere die Kontrolle von Störungen. Am Beispiel einer Skala zur

Messung von Konservativismus erörtern *Cloud* und *Vaughan* Möglichkeiten zur Kontrolle von "acquiescence" (Tendenz zur Zustimmung), der Neigung bei vielen Befragten, Sätzen ungeachtet des Inhaltes zuzustimmen (61). Die beiden Autoren versuchen für die von ihnen benutzte Skala nachzuweisen, daß ein bloßes Alternieren von Sätzen, die Zustimmung und Ablehnung erfordern, bereits ausreicht, um diesen Zustimmungseffekt zu neutralisieren. Da diese Technik bereits Anfang der fünfziger Jahre entwickelt und begründet wurde, ist die Notwendigkeit dieser Publikation nicht offensichtlich; vielleicht wurde der Artikel vor allem deshalb veröffentlicht, weil die Autoren Neuseeländer sind. Mit der Kontrolle von Zustimmungseffekten befaßt sich auch ein Bericht von *Leslie G. Carr*, in dem gleichfalls die Wirkung alternativer Formulierungen von Skalen-Aussagen dargestellt wird (62). Die Stärke einer Zustimmungseigung ungeachtet des Inhalts einer Aussage wird hier für die seit 1956 in den USA viel benutzte "Anomie-Skala" von *Leo Srole* (ob überhaupt die Skala mit Anomie viel zu tun hat, ist eine andere Frage) nachgewiesen. Auch hier ist angesichts vieler ähnlicher Untersuchungsberichte aus früheren Jahren nicht ganz einsichtig, warum nun auch dieser noch erscheinen mußte – es sei denn, die Redakteure wollten angesichts der häufigen Verwendung der Skala von *Srole* die Forschungspraktiker daran erinnern, daß die Anomie-Skala ungeachtet aller Problematisierung in der methodologischen Literatur zu routinehaft angewandt wird.

Schließlich wird noch ein der Zustimmungseigung verwandter Störfaktor bei Skalenmessungen, "social desirability" oder soziale Akzeptierbarkeit, bei Erhebungsgegenständen wie Religiosität und Glücksgefühl untersucht (63). Die Autoren schlagen als Ergebnis vor, von jetzt ab zwischen zwei Störfaktoren zu unterscheiden, "trait desirability" (Wünschbarkeit der Eigenschaft) und "need for social approval" (Bedürfnis nach sozialer Annehmbarkeit), deren unterschiedliche Beziehung zu sozialen Merkmalen bisher durch ihre Zusammenfassung zu sozialer Akzeptierbarkeit verdeckt worden sei. Dies ist die Art von Untersuchung und Bericht darüber, die früher für die methodologische Literatur in ihren routinehaften Ausdrucksformen kennzeichnend war: Man nehme sich eine bekannte Variable und weise nach, daß sie in Wirklichkeit mehr als eine Dimension repräsentiere bzw. anders als bisher behauptet mit demographischen Merkmalen korreliere. Für die praktische Forschung und die Weiterentwicklung der Methodologie bleiben solche Berichte ebenso folgenlos, wie die heute in der Literatur wuchernden Varianten statistischer Maßzahlen.

Verfahrensweisen, die den Rechenaufwand bei der Prüfung von Skalen vermindern sollen, sind ein weiteres Standardthema der Literatur zu Skalierung. Die Forschungspraktiker scheuen wegen des Rechenaufwandes nicht selten Prüfungen der Verlässlichkeit und Gültigkeit, sowie der Dimensionalität von Meßinstrumenten, sodaß solche Verfahrenstechniken durchaus einen Nutzen haben könnten. Das von *George W. Bohrnstedt* vorgestellte Verfahren soll die Berechnung sowohl der Verlässlichkeit als auch der Gültigkeit erleichtern, und dies für die Eigenschaft einer Skala, zwischen Fällen zu diskriminieren ("discriminant validity") sowie für die weitere Eigenschaft der Skalenwerte, mit äußeren Kriterien übereinzustimmen ("criterion-related validity") (64). Die Arbeitersparnis als Folge dieses Verfahrens scheint sich aber lediglich auf das Vermeiden einiger einfacher Rechenoperationen zu beschränken, was bei der Zugänglichkeit moderner Computer

als Argument etwas schwach scheint. Dagegen beruht die Empfehlung von *Roland J. Chilton* für ein bestimmtes Verfahren zur Prüfung von Guttman-Skalen auf einem Vergleich verschiedener Alternativen zur Berechnung des "Koeffizienten der Reproduzierbarkeit" (coefficient of reproducibility), der allgemein üblichen Prüfung von Guttman-Skalen auf Eindimensionalität (65). Das damals von Guttman selbst vorgeschlagene Verfahren war zwar plausibel, enthielt aber kein statistisches Kriterium zur Bestimmung der Signifikanz eines Ergebnisses; inzwischen liegt ein solches Verfahren vor, mit dem ein Vergleich zwischen Erwartungs- und Beobachtungswerten möglich wird ("chance reproducibility"). Die Prüfung an hypothetischen und empirischen Daten ergibt die Überlegenheit dieser Technik, für die zudem auch ein Computer-Programm verfügbar ist – womit erklärungsbedürftig wird, warum in der Praxis der Forschung immer noch Verfahren benutzt werden, die in Lehrbüchern als problematisch gekennzeichnet sind. Zudem wird in einem weiteren Untersuchungsbericht noch einmal nachgewiesen, wie groß bei den meisten der bisher benutzten Verfahren zur Prüfung der Reproduzierbarkeit die Chance ist, rein zufällig angemessene Werte zu erhalten (66).

Die deutschsprachigen Beiträge haben auch beim Thema Skalierung den gleichen, von der englischsprachigen Literatur abweichenden Charakter, wie er bereits zum Thema Pfadanalyse beobachtet wurde. Beide Aufsätze haben wiederum den Charakter eines Beitrags zu einem Lehrbuch (67). Dies ist nicht als Tadel gemeint, denn *Kurt Holm* legt mit diesen Aufsätzen eine sehr gut lesbare und informierte Übersicht der verschiedenen Verfahren zur Prüfung von Gültigkeit und Verlässlichkeit vor (68). Es wird an diesem Beispiel jedoch die andere Beziehung zwischen den Publikationsformen Zeitschrift, Reader und Lehrbuch außerhalb und innerhalb der Vereinigten Staaten deutlich.

In den Aufsätzen zu methodischen Problemen der Soziometrie geht es nur noch um Fragen der Analyse, wobei die Analyse der Cliques-Struktur besonderes Interesse fand. In einem Vergleich von fünf Verfahrensweisen kann nachgewiesen werden, daß die Befunde über Art der Cliques-Struktur mit abhängen von der Verfahrensweise; dabei ergibt sich jedoch eine eindeutige Präferenzordnung, in der Faktorenanalysen erst an letzter Stelle kommen (69). Als Alternative zu Verfahrensweisen, die einen geschlossenen Rechengang verwenden, wird eine iterative Technik vorgestellt, in der die Unterschiede zwischen verschiedenen Arten von Cliques deutlich werden (70). Eine generelle Darstellung der Struktur interpersoneller Beziehungen über die Analyse von Cliques hinaus – für die sich allerdings leichter mathematische Modelle entwickeln lassen, vor allem bei einer Beschränkung auf Paarwahlen! – wird mit der Absicht vorgestellt, die verschiedenen Cliques-Analysen als spezielle Fälle zu integrieren (71). Ein besonderer Vorzug dieses Verfahrens dürfte es sein, daß Erwartungswerte berechnet werden können. Ebenfalls mit der Anwendung wahrscheinlichkeitstheoretischer Kalküle arbeitet ein Verfahren, daß sich wieder auf gegenseitig erwiderte Sentiments konzentriert (72). Die Besonderheit dieses Beitrags ist die Verbindung von Soziometrie mit der Rollen-"Theorie", indem hier der Personensatz (im Verständnis von *Robert K. Merton*) soziometrisch abgebildet wird. Insgesamt ergibt sich aus der Betrachtung von methodologischen Aufsätzen zur Soziometrie der Eindruck, daß sie wesentlich von Spezialisten für Spezialisten verfaßt sind, daß die Aufsätze jeweils eine unmittelbar einsehbare Weiterentwicklung des Instrumentariums bringen und

daß es sich hier um ein Arbeitsgebiet handelt, das sich kontinuierlich weiter entwickelt. Noch Anfang der sechziger Jahre konnte demgegenüber der Eindruck gewonnen werden, daß die Soziometrie in zahlreiche Spezialitäten bis hin zu esoterisch-inhaltslosen Sekten zerfasere (73); hier hat offensichtlich nicht zuletzt die Anwendung mathematischer Verfahren eine Wende bewirkt.

Ende der fünfziger und Anfang der sechziger Jahre gab es eine lebhaftige Kontroverse über die legitimen Anwendungsmöglichkeiten sogenannter ökologischer Daten und generell über Aggregatdaten-Analysen. Inzwischen ist nicht mehr kontrovers, daß bei jeder Verwendung von Aggregatdaten als ein stellvertretendes Messen von Eigenschaften auf einer niedrigeren Ebene der Aggregation die Gefahr eines Fehlschlusses gegeben ist, daß aber generell durch die Verwendung von Aggregatdaten die Aussagebreite der Forschung verstärkt wird (74). Nach Klärung dieser Grundsatzfragen kann sich die methodologische Literatur auf das "wie" der Verwendung von Aggregatdaten konzentrieren. So klärt *Hammond* die besonderen Bedingungen, unter denen keine Rückschlüsse von Aggregatdaten auf Individualdaten mehr möglich sind: Wenn die unabhängige Variable für die Gruppierungen einen Kontexteffekt hat (75). Auf der Grundlage von empirischen Daten zeigt *Iversen*, wie aus aggregierten doch wieder Individualwerte rückgewonnen werden können (76). Die Aggregatdaten haben jedoch nicht nur und inzwischen nicht einmal in erster Linie Bedeutung als eigene Meßebeine für sich, sondern in Kombination mit Individualdaten. In einem der Beiträge wird aufgezeigt, wie eine Aggregation der Fälle nach der unabhängigen Variablen in einer Verbesserung der Erklärung resultiert, während eine Gruppierung nach der abhängigen Variablen zu Artefakten führt (77). Auf der Grundlage empirischer Daten wird in einem anderen Beitrag der Nachweis geführt, daß Kontexterklärungen von unterschiedlichen Reaktionen bei Individuen mit den gleichen Merkmalen – hier handelt es sich um den vorgeblichen Effekt des Schulmilieus auf die Motivation von Schülern – ebenso zu Fehlschlüssen führen können, wie andere Arten von Scheinkorrelationen (78). Dies ist eine nützliche Erinnerung daran, daß wir noch nicht über eine Fehlerlehre für Mehrebenen-Analysen verfügen. Insgesamt ist jedoch Umfang und Gehalt der Artikel, die sich mit Aggregatdaten-Analysen und Mehrebenenanalysen befassen, sehr enttäuschend.

Die für die allgemeine Methodologie wichtigsten Aufsätze lassen sich um das Thema "undermeasurement" gruppieren; "Meßniveau" wäre für die einseitige englischsprachige Formulierung ein besserer Ausdruck im Deutschen. Bekanntlich ist die große Mehrzahl der statistischen Verfahren für Daten ausgebildet, die mindestens dem Niveau des Intervallmessens entsprechen, während die große Mehrzahl aller Daten in der Sozialforschung ordinales Meßniveau nicht überschreitet – höchstens nicht selten ein Zwischenniveau erreicht, das man als "ordinale Intervalle" bezeichnen kann (79). Was macht es aus, wenn Intervallmaße auf Daten angewandt werden, die ordinales Messen als Niveau nicht überschreiten? ist die Fragestellung in dieser Kontroverse zum Thema "undermeasurement". Einerseits kann argumentiert werden, daß eine Beschränkung auf Maßzahlen für ordinales Messen zu einem Verlust an Aussagekraft führt, weil eben diese Maßzahlen wegen der geringeren Unterstellungen des Algorithmus auch nur weniger aussagen können; andererseits wird durch Maßzahlen für Intervallmaße, die auf ordinale Daten angewandt werden, die Gefahr von artifiziellen Ergebnissen her-

vorgeworfen. Diese beiden Standpunkte – die Befürwortung einer Anpassung der Maßzahlen an das niedrigere Meßniveau, oder die Verwendung der aussagekräftigsten Maßzahlen auch bei einem niedrigeren Meßniveau der Daten – haben sich während der sechziger Jahren in den Vereinigten Staaten zu veritablen Schulen verfestigt, die die Etikette “schwache Messer” (weak measurement) und “starke Statistiker” (strong statistics) erhielten (80). Entgegen den Plausibilitätsvorstellungen des alltäglichen Denkens ist es nicht möglich, sich aus der Affaire zu ziehen, indem man sich aus Gründen der Vorsicht für Maßzahlen nur des ordinalen Meßniveaus entscheidet: Es ist eine Entscheidung notwendig, ob man einen Fehler I oder Fehler II (irrtümliche Annahme eines in Wirklichkeit nicht belegten Ergebnisses bzw. irrtümliches Übersehen eines in Wirklichkeit doch gegebenen Befundes) für wahrscheinlicher und folgenreicher hält. Es ist mithin sowohl eine meßtechnische als auch eine inhaltliche Aussage notwendig.

Inzwischen dürfte sich jedoch die Kontroverse über das Meßniveau – die bisher als ein allgemeines Thema außerhalb der USA noch nicht als solches rezipiert wurde – vom Schulentreit zu einer Spezifizierung der Bedingungen weiter entwickeln, unter denen der Standpunkt der “schwachen Messer” und die Umstände, unter denen der Haltung der “starken Statistiker” gerechtfertigter ist (81). Allgemein überwiegen die Autoren, die eine Verletzung der Annahmen über Meßniveau für sehr viele Maßzahlen nicht für sehr gravierend halten und demgegenüber den Gewinn an Aussagekraft für das entscheidende Argument ansehen (82). Tatsächlich sind Verletzungen von Annahmen über Meßniveau bei solchen Maßzahlen wie Mittelwerten oder Korrelationsmaßen nur unter bestimmten Bedingungen gefährlich, die es jetzt genauer als bisher zu spezifizieren gilt. Weitgehende Übereinstimmung dürfte inzwischen darüber erzielt worden sein, daß bei ordinalen Daten, bei denen Intervall- oder Ratio-Maßzahlen angewandt werden, die Gefahr von Fehlschlüssen größer ist, wenn als Grundlage der Berechnung tabellarisch aufbereitetes Material verwandt wird. Dies ist insbesondere bei Berechnungen aufgrund der beliebten 2 x 2 Tabellen zu beachten (83). Überhaupt muß die Neigung insbesondere vieler amerikanischer Sozialforscher kritisch beurteilt werden, ihre Daten vor der Anwendung von Maßzahlen stark zu vergrößern; nicht wenige “Ergebnisse” in der Standardliteratur dürften einer Neuanalyse nicht standhalten. Allerdings gibt es erst für bestimmte Bedingungen Maßzahlen, welche den Effekt einer Zusammenfassung von Einzelwerten zu Gruppen ausdrücken bzw. die angemessenen Intervalle und Zahl der Gruppen angeben – und diese Maßzahlen sind meines Wissens nicht Teil der Ausbildung in Statistik.

Die größte Bedeutung dieser Kontroverse über Meßniveaus dürfte darin bestehen, daß nun von der Sozialforschung aus die sehr ungenügende Fehlerlehre für statistische Maßzahlen weiterentwickelt wird. Darauf deuten solche Beiträge hin, in denen erörtert wird, wie man mit fehlerhaften Mehrfachindikatoren für die einzelnen Faktoren in einer Pfadanalyse doch rechnen kann (84); oder bis zu welchem Grade bei Unvollständigkeit der Daten dennoch eine Matrizenrechnung möglich bleibt (85). Besonderes Interesse fanden auch die Umstände, unter denen Korrelationsmaße durch Verletzung der Unterstellung über Verteilungsform bzw. der latenten Faktorenstruktur beeinflußt werden (86). Da die Annahmen über den Charakter der Daten in der Pfadanalyse besonders anspruchsvoll sind, ist verständlich, daß hierauf relativ häufig eingegangen wird, auch mit dem Ziel, durch

Vorschläge von Datentransformationen und verwandte Manipulationen selbst bei zunächst ungeeigneten Daten dennoch Berechnungen vornehmen zu können (87). Sollte diese Tendenz in der Fehlerlehre fortbestehen, so wird zusätzlich zur Klärung der Robustheit von Maßzahlen gegenüber Verletzungen ihrer Kalküle bzw. Annahmen eine sehr verästelte Kasuistik der "Schönung" von Daten vor ihrer Nutzung für Berechnungen entsteht. Dies wird dann selbst ein Sozialforscher nicht mehr übersehen können, der sich auf Analyse spezialisiert, womit aus dieser Entwicklung die Ausdifferenzierung einer weiteren Spezialität in der Methodologie folgen könnte.

Die Probleme der Datensammlung treten während des Beobachtungszeitraums gegenüber den bisher erwähnten Themen zurück bis zur quantitativen Bedeutungslosigkeit. Eine der wenigen Ausnahmen in diesen amerikanischen Zeitschriften ist eine Untersuchung über die Verlässlichkeit von Verschlüsselungen von Interview-Daten – also selbst dies keine Behandlung der Datensammlung i.e.S. (88). Die bedeutendsten Beiträge – nicht Methodologie im abstrakten Sinne, sondern auf der Grundlage empirischer Forschung – erscheinen in ausländischen Zeitschriften, darunter drei wichtige Beiträge zur Frage im Interview als Mittel der Sozialforschung (89). Während in amerikanischen Quellen Probleme der Inhaltsanalyse in Verbindung mit Computerprogrammen dargestellt werden, ist die inhaltliche Auseinandersetzung mit Inhaltsanalyse wieder nur in einer deutschsprachigen Zeitschrift zu finden (90). So ergibt sich für die Datensammlung ein im Vergleich zur Datenanalyse spiegelverkehrtes Bild der methodologischen Literatur in den USA und in der Bundesrepublik: Während hier die Zeitschriften ein wichtiges Kommunikationsmittel sind, gewinnt für die Methodologie der Datensammlung in den USA das Buch eine zunehmende Bedeutung (91).

Eine teilweise Ausnahme von dieser übermäßigen Konzentration auf die Datenanalyse – übermäßig, denn in den Auseinandersetzungen über Meßniveaus wird ja doch die Qualität der Datensammlung thematisiert, allerdings auf ein technisches Problem verkürzt – sind Fragen der Versuchsanordnungen bzw. der Forschungspläne ("design"). In einer Übersicht verschiedener bekannterer Untersuchungen wird für eine Kombination qualitativer und quantitativer Verfahren in der Erhebung plädiert (92). Ebenfalls auf der Grundlage vorliegender Monographien über Untersuchungen wird erörtert, welche Techniken bei der teilnehmenden Beobachtung als Verfahren der Datensammlung den Gegenstand möglichst störungsfrei abbilden (93). Größer ist jedoch die Zahl der Aufsätze, die sich mit speziellen Fragen der experimentellen Versuchsanordnungen befassen. Dazu gehören solche Themen wie die Berücksichtigung der Interpretation des Stimulus durch die Versuchsperson, die Kontrolle externer Einflüsse und die Rollenprobleme verdeckter Helfer des Experimentators (94). Grundsatzfragen von Versuchsanordnungen werden allerdings nicht aufgeworfen: Es handelt sich um zusätzliche Vorschläge zu bekannten Problemstellungen.

Soziale Mobilität ist während des Beobachtungszeitraums ein inhaltliches Thema gewesen, das zu interessanten methodischen Erörterungen anregte (95). Sie folgen aus dem meist unbefriedigenden Zustand der Daten, insbesondere für die Untersuchung der Intergenerations-Mobilität. Ein Beispiel für die Art der methodischen Fragen dieses Sachgebiets ist eine Untersuchung der Fehlschlüsse, die sich aus der üblichen Messung von Mobilität als Differenz zwischen dem sozia-

len Status des Vaters und dem des Sohnes ergibt; vorzuziehen soll eine Abbildung des väterlichen Status als Individualmerkmal des Sohnes sein. Ein weiteres Thema, das häufiger zu methodologischen Erörterungen anregte, ist die Erklärung von Einstellungen und deren Wirkungen. Auch hier geht es durchweg um technische Verbesserung – wie etwa in einer Untersuchung dreier Berechnungsverfahren, mit denen die Abhängigkeit von Verhalten als Produkt von Einstellung und Situation ausgedrückt werden soll (96).

All dies erschien während eines Zeitraums, der für die intellektuelle Öffentlichkeit durch Grundsatzdiskussionen gekennzeichnet schien. Nun war selbstverständlich nicht zu erwarten, daß Literatur für eine Fachöffentlichkeit getreu die Themen widerspiegelt, die in einer allgemeineren Öffentlichkeit dominierten. Aber dies war ja kein normaler Zeitraum: Er schien durch eine Umwertung vieler Werte gekennzeichnet, und die Soziologen waren durch recht bekannte Personen an dieser Diskussion beteiligt. Auf Soziologentagen wurden außerwissenschaftliche Festlegungen zu soziologischen Weltanschauungsparteien hochstilisiert, wobei diese Parteien sich durch ein unterschiedliches Methodenverständnis voneinander abgrenzen sollten. Bei dieser Auswertung war erwartet worden, daß sich dieses Klima in einer Bevorzugung von Grundsatzfragen des wissenschaftlichen Ansatzes ausdrücken würde, oder daß zumindest auch bei der Behandlung eher forschungstechnischer Fragen diese in ihren Annahmen öfters "hinterfragt" würden. Davon kann nun nach Durchsicht des Schrifttums überhaupt keine Rede sein. Es wurde nur ein Aufsatz gefunden, der diesen Grundsatzfragen nahe kam: Eine Erörterung des "Reduktionismus", d.h. der Zurückführung soziologischer Aussagen auf Sätze der Psychologie; diese Grundsatzfrage wurde jedoch in ihrer Behandlung zu einer Frage der Machbarkeit (97). Auch eine Behandlung der Beziehung zwischen Begriffen und Meßoperationen bei Typologien wurde in der Ausführung zu einer Frage der Ersetzung von bisherigen Taxonomien von Typologien durch eine andere Art des Ordnen von Typen und Skalen (98).

4.

Bis hierher wurden in der Übersicht die Beiträge, die im "Public Opinion Quarterly" erschienen, nicht berücksichtigt, obgleich diese Zeitschrift vom Volumen methodologischer Aufsätze her die wichtigste Quelle ist. Es erweist sich jedoch, daß sich die Leserschaft des "POQ" mit derjenigen der bisher ausgewerteten Fachzeitschriften für Soziologie weniger überschneidet, als dies für diese untereinander gilt. Die Thematik der Aufsätze ist anders, mit dem eindeutigen Schwerpunkt auf Problemen der Umfrageforschung; diese waren demgegenüber in der bisher behandelten Literatur zweitrangig.

Wegen der relativ größeren Einheitlichkeit der Art der Aufsätze, die sich an eine andere und in ihren methodischen Interessen einheitlichere Klientel richten, erschien es hier eher als bei dem vorher erörterten Material vertretbar, auch zu zählen. Aus der folgenden Übersicht sollte deutlich werden, daß eine getrennte Besprechung der Aufsätze in dieser Zeitschrift statt ihrer Vermischung mit den anderen Publikationen sinnvoll war (99).

Verteilung der Themen im "Public Opinion Quarterly" im Zeitablauf 1968-1974 für Aufsätze mit methodologischem Schwerpunkt

Thema	Summe	1968	1969	1970	1971	1972	1973	1974
1 Stichproben	5	1	0	1	2	0	0	1
1a Speziell: Ausschöpfung	19	1	5	2	1	3	1	6
2a Interviews	9	2	1	3	1	2	0	0
2b Schriftliche Befragungen	12	2	1	3	1	3	2	0
2c Administration der Feldarbeit	5	1	1	0	0	1	1	1
2d Der Befragte als Störung	16	1	2	2	5	2	1	3
3a Panelfragen	7	1	0	1	1	0	1	3
3b Experimente und Sonder- techniken	11	0	4	2	2	1	1	1
4a Analysefragen	8	0	2	1	0	1	2	2
4b Sekundäranalysen und Datenbanken	4	1	2	1	0	0	0	0
4c Skalierung	5	0	0	0	1	1	3	0
4d Verschlüsselungsfragen	3	2	0	0	0	1	0	0
Summe	104	12	18	16	14	15	12	17

Neun von den insgesamt für diese Auswertung bestimmten 115 Aufsätzen konnten in das vorstehende Klassifikationsschema nicht eingeordnet werden, und sie waren zugleich in sich zu heterogenen, um für sie eine eigene Kategorie abzugrenzen. Es handelte sich hier beispielsweise um die Ankündigung eines neuen Service durch NORC, oder um einen neuen Typ von Umfragen durch das Gallup Institut ("opinion referendum") (100). Es wird die Nützlichkeit der Zensus-Daten für Umfrageforscher erörtert (101), oder die Aussagekraft von Vorformen der Umfrageforschung in der Renaissance (102). Methodologisch-begriffliche Aufsätze, wie die Gültigkeit einer Formulierung des Begriffs "Meinungsführerschaft" (103), waren ebenfalls zu selten, um hierfür eine eigene Kategorie abzugrenzen.

Selbst zum heutigen Zeitpunkt sind wichtige Grundsatzfragen der Stichprobenverfahren für Umfragen ungeklärt (104). So wird die Formel zur Berechnung des Stichprobenfehlers der Tatsache nicht gerecht, daß im Gegensatz zum Modell in einer Stichprobe hunderte von Merkmalen gleichzeitig erhoben werden; dadurch wird die Aussagekraft der Signifikanzberechnungen von Korrelationsmaßen auf unbekannte Weise beeinträchtigt. Oder es ist unbekannt, welcher Fehler durch mehrstufige Identifizierung der zu beobachtenden Einheiten dem heute tatsächlich berechneten Fehler hinzugefügt wird; die Praxis berücksichtigt diesen Umstand durch einen nur pragmatisch zu rechtfertigenden Korrekturfaktor $\sqrt{2}$. In der Zeitschrift werden demgegenüber durchweg praktische Fragen der Stichproben behandelt, wie die Angemessenheit von Telefonbüchern als Basis für Stichproben, oder die Möglichkeit der Nutzung von Telefonnummern zur Auswahl von Fällen (105). Dies sind ziemlich kulturspezifische Themen für ein Land, das keine Einwohnerverzeichnisse hat und in dem das Telefon allgemeiner verbreitet ist als in irgendeinem anderen Land. Von der Situation der Praxis her besteht

auch kein aktueller Anlaß, die verschiedenen offenen Fragen der Anwendung von Stichprobenmodellen auf menschliche Bevölkerungen zu behandeln. Schwierigkeiten ergeben sich aus anderen Gründen, insbesondere aus der unvollkommenen Ausschöpfung der Stichproben.

Wie die Auszählung ergab, wird kein einzelnes forschungstechnisches Thema in POQ so häufig behandelt wie die unvollkommene Ausschöpfung. In den letzten 10 Jahren scheint die Kooperationsbereitschaft bei mündlichen und schriftlichen Befragungen abzunehmen, sodaß die Verweigerung unter den Gründen für Ausfälle an Bedeutung zunahm. Diese für die Aussagekraft von Umfragen sehr ernste Entwicklung spiegelt sich in den Themen wieder, wenn die Wirksamkeit der Kombination von schriftlicher Befragung und persönlichen Besuchen bei nicht-Antworten untersucht oder die Wirkung von zusätzlichen Zahlungen geprüft wird – bis hin zu so technischen Einzelheiten wie die Gestaltung der Rückantwort-Umschläge zur Erhöhung der Ausschöpfung (106). Allgemein ist die Behandlung des Themas Ausschöpfung verbunden mit praktischen Vorschlägen, wie die Verluste (sampling losses) verringert werden könnten (107). Da hier verfahrenstechnisch nur begrenzte Erfolge möglich sind, bleiben als wichtigstes Korrektiv doch wieder die seit über 20 Jahren umstrittenen nachträglichen Gewichtungen aufgrund von Kontrollen der nicht-erreichten Fälle (108). Insgesamt kann diese Gruppe von Aufsätzen als eine laufende Anpassung an veränderte Bedingungen in der Praxis qualifiziert werden.

Auffällig ist bei der Behandlung der Datensammlung die große Aufmerksamkeit für Probleme der schriftlichen Befragung. Dies ist die wahrscheinlich bedeutendste Veränderung in den Schwerpunkten der Zeitschriftenliteratur zur Umfrageforschung. In den fünfziger Jahren dominierten nach Zahl und Wichtigkeit Artikel über Probleme des (persönlichen) Interviews: Da wurden die Wirkung von Frageformen und -anordnungen untersucht, die Bedingungen des Interviews und vor allem das Verhalten des Interviewers (109). Der Interviewer wurde als eine der wichtigsten Fehlerquellen des gesamten Umfrageprozesses angesehen; und wenn berücksichtigt wurde, auf welche Fehlerquellen der Forscher noch relativ den größten Einfluß ausüben kann, als wichtigste Fehlerquelle. Nun hat sich nach dem "Folklore" unter Umfragespezialisten die Qualität der Interviewer seither nicht verbessert – und dennoch sind nicht nur ihrer geringen Zahl nach die Aufsätze über (mündliches) Interviewen überraschend beschränkt. Gelegentlich werden die "klassischen" Themen der früheren Literatur wieder bearbeitet, wie die angemessene Formulierung der Frage in strukturierten Fragebögen, oder die Wirkung sozialer Distanz zwischen Befragtem und Interviewer, sowie die mündliche Befragung als ein Rollenverhalten (110). Häufiger sind jedoch als Thema sehr spezielle Fragen, wie die zweckmäßigste Tageszeit für Interviews mit Erwerbstätigen oder Probleme bei Befragungen alter Menschen; man erfährt über beides nicht wesentlich Neues (111). Dieser Zustand der Literatur kann spekulativ auf zweifache Weise gedeutet werden: Während der früheren Konzentration auf dieses Thema sind die wesentlichsten Fragen soweit geklärt und die Organisation der Befunde ist soweit abgeschlossen, wie es das Orientierungsbedürfnis der Praktiker erfordert; und/oder die sinkende Kooperativität der Befragten und aus Kostengründen das Ausweichen auf Sonderformen der Befragung lassen diese Aspekte des Umfrageprozesses dringlicher werden. Diese Charakterisierung gilt übrigens

für die Bundesrepublik nicht, da hier über Methodenfragen des Interviews beachtliche Forschungsarbeiten gerade während des Beobachtungszeitraums vorgelegt wurden (112).

Die schriftliche Befragung in ihren verschiedenen Varianten – postalisch, selbst auszufüllende Fragebögen allein oder in Gruppensituation administriert – hatte lange in der kommerziellen Forschung keine mit der mündlichen Befragung auch nur in etwa vergleichbare Bedeutung. Wegen der Kostengünstigkeit und vor allem wegen der höheren Bildung der angesprochenen Population wurden schriftliche Befragungen von Sozialforschern an Hochschulen bevorzugt. Zudem schienen die methodologischen und auch im engeren Sinne forschungstechnischen Fragen schriftlicher Befragungen hinreichend geklärt. Diese Einschätzung wird indirekt durch die betreffenden Veröffentlichungen, trotz ihrer verblüffenden Häufigkeit, bestätigt. Nur selten werden allgemeinere Problemstellungen behandelt, wie etwa die Wirkung unterschiedlicher Formen der Administration von Fragebögen, und dabei werden auch keine wesentlichen neuen Befunde ermittelt (113). Das beherrschende Thema sind die Anreize zur besseren Mitarbeit der Befragten bei postalischen Erhebungen, die in erster Linie durch Variationen wie Ankündigungsbriefe, durch persönlichere Ansprachen und weitergehende Vorverschlüsselungen in ihrer Ausschöpfungsquote und der Qualität der Antworten in ihrer Wirksamkeit erhöht werden sollen (114). Es hat jedoch den Anschein, als ob längerfristig ein Mittel nicht zu vermeiden sein wird, gegen das sich die privatwirtschaftlichen Institute sträuben: Materielle Belohnungen für gut beantwortete Fragebögen (115). Bis dahin wird sich eben die geringer gewordene Motivation von Befragten zur Mitarbeit, die sich bei mündlichen Befragungen bereits als gegenwärtiges Hauptproblem erwies, bei allgemeinen postalischen Befragungen in zu niedrigen Ausschöpfungsquoten ausdrücken (116).

Es war charakteristisch für die Zeitschriftenliteratur zur Methodologie der Umfrageforschung in den fünfziger Jahren, daß als wichtigste Möglichkeit zur Verbesserung der Erhebungen die Kontrolle des Interviewers angesehen wurde. Einzelprobleme der Feldarbeit wurden entsprechend häufiger und informativ behandelt, während jetzt die Auswertung der Zeitschriftenliteratur nur wenige und nicht besonders beachtliche Artikel ergab (117). Demgegenüber haben Untersuchungen über den Befragten als Teil des Erhebungsprozesses gegenüber früheren Zeiträumen in einem Maß zugenommen, das schon als explosionsartig bezeichnet werden kann. Auch hier werden verschiedene Einzelfragen aufgegriffen (118), aber zwei Problemstellungen dominieren: (a) die Verlässlichkeit der Antworten und (b) der Befragte als Ursache mangelnder Gültigkeit von Befunden. In der Behandlung von Verlässlichkeit geht es verschiedentlich um Definitions- und Meßprobleme (119), und die gelegentlichen empirischen Untersuchungen sind nur von Routine-Interesse (120).

Das gilt auch für einige Untersuchungen zum Thema Gültigkeit, in denen Antworten mit objektiven Unterlagen verglichen wurden. Wird die Auswertung nicht von inhaltsreichen Hypothesen geleitet, dann gibt es ein Nebeneinander von Daten mit hoher bis zureichender Gültigkeit und solchen mit weniger befriedigender Gültigkeit, dessen Verallgemeinerung über den mitgeteilten Fall hinaus schwierig bis unmöglich wird (121). Dies gilt auch für die Erhebungen bei sehr speziellen Gruppen, wie den von psychotropen Drogen Abhängigen, wenn eben der Bezug

zu allgemeineren Determinanten als denen, welche die Besonderheit der Gruppe konstituieren, unklar bleibt (122).

In den USA ist die Voraussage der Wahlbeteiligung schwieriger als in der Bundesrepublik, weil zusätzlich zur Wahrscheinlichkeit des Wählens auch noch die Chance zu bestimmen ist, daß sich ein Befragter in die Wählerlisten einträgt; oft wird die Wahlbeteiligung um 10-15% überschätzt. Ein Teil dieser Übertreibung soll nach einer der Untersuchungen darauf zurückzuführen sein, daß das Interview den Befragten verändert, indem es ihn zum Wählen anregt – eine Beobachtung, die als "Stimulierungs-Hypothese" vom Autor verallgemeinert wird (123). Eine andere Verallgemeinerung mit der gleichen Richtung, daß nämlich in vielen der Gültigkeitsprüfungen das Ausmaß der Ungültigkeit übertrieben dargestellt werde, ergibt sich aus einer Untersuchung von Gewerkschaftsfunktionären. Eine Befragung erbrachte, daß von 64% aller Befragten die Zahl der besuchten Schulungskurse unrichtig angegeben wurde, gewöhnlich in Richtung auf die erwartete Antwort einer möglichst hohen Teilnahme; das Ausmaß der Abweichung blieb jedoch gering, sodaß für praktische Zwecke die Daten als dennoch brauchbar behandelt werden konnten (124). Die aus beiden Untersuchungen abzuleitende Schlußfolgerung wäre ein geringer Anteil von Befragten, der als "Wahrheitsrisiko" eingestuft werden müßte. Selbst wenn dem zuzustimmen wäre – was gerechtfertigt sein dürfte – ist damit wenig über die Problematik der Gültigkeit von Antworten ausgesagt; in erster Linie ist damit nur ein verbreitetes Vorurteil der Laienöffentlichkeit relativiert.

Es ist in der Umfrageforschung verbreitet, nur eine Person pro Haushalt zu interviewen, deren Antworten aber teilweise als stellvertretende Aussage über die anderen Personen des Haushaltes zu verwerten. Der Vergleich von Antworten eines Ehepaars, das getrennt über die gleichen Dinge befragt wurde, ist einmal ein Test der Zulässigkeit einer solchen Stellvertreterschaft und liefert je nach Frage auch ein Indiz für zweifelhafte Gültigkeit. In einer Befragung von 179 Ehepaaren wurde u.a. die Übereinstimmung der Angaben für Daten wie Einkommen und Aussagen über Verantwortlichkeit für Kindererziehung geprüft (125). Wieder ergab sich, daß die Übereinstimmung bei einem "harten" Merkmal wie Einkommen – "hart" definiert als einen Sachverhalt, der unabhängig von der Person eindeutig zu bestimmen ist – recht hoch war: 60,3% völliger Konsens. Dagegen gaben von ebendiesen 60,3% der Befragten 46,3% unterschiedliche Antworten zur Frage nach der Aufgabenteilung bei der Kindererziehung. Tendenziell schrieb sich dabei jeder der Ehegatten jeweils ein höheres Maß an Verantwortlichkeit zu, als der Partner angab. Das Ergebnis kann als Beleg für das Ausmaß an Perzeptionsverzerrungen bei solchen Sachverhalten genommen werden, die bei den Befragten Raum für unterschiedliche Interpretationen belassen. Diese Aussage mag intellektuell trivial sein, ist aber mit den hier möglichen Angaben über die Größenordnung eventueller Verzerrungen für die Praxis wichtig, in der die unterschiedliche Härte von Daten nicht zureichend beachtet wird – wie etwa bei vielen Meßinstrumenten.

Personen mit niedrigem sozialen Status und/oder einer randseitigen Position sollen nach dem unter Umfrageforschern verbreiteten Folklore besonders häufig ungültige Antworten geben. In einer Erhebung des "National Opinion Research Center" (NORC) wurden die Angaben von Neger-Müttern, die Wohlfahrtsempfänger waren, für fünf Fragen mit amtlichen Unterlagen verglichen (126). Dies ist eine Übersicht über die Fehlerhaftigkeit der Antworten:

Eintragung in das Wahlregister	18%	Fehler
Teilnahme an der Wahl zum Präsidenten	18%	”
Blieb das eigene Kind jemals sitzen?	22%	”
Wurde das eigene Kind im letzten Jahr nicht versetzt?	37%	”
Zahlungen der Wohlfahrtsbehörden	2%	”

Wichtiger als die Angaben über die Höhe der Gültigkeit von Angaben zu verschiedenen Themen – die allerdings beim Kriterium “Zahlungen” überraschen – sind die Korrelate von Ungültigkeit. Zunächst wird wieder einmal belegt, was bisher zu belegen für Meinungen über Umfrageforschung offensichtlich keine Wirkung hat: daß es Befragte mit durchgehend ungültigen Angaben nur als Ausnahmefälle gibt; die Fehlerhaftigkeit der Antworten war untereinander überwiegend nicht korreliert (127). Die soziale Distanz zwischen Interviewer und Befragtem hatte wenig Einfluß, wogegen ein besonders guter Rapport zwischen den Partnern im Interview mit geringerer Gültigkeit einherging. Beides widerspricht verbreiteten Annahmen über das Interview, ist aber bereits verschiedentlich nachgewiesen worden (128).

Generell waren die Korrelationen zwischen der fehlenden Gültigkeit von Antworten zu einzelnen Fragen und einzelnen Variablen, die nach dem Folklore die Gültigkeit von Angaben verringern, nicht sehr bedeutend – und doch hat das Folklore nicht einfach unrecht, sondern stellt eine unzulässige Vereinfachung dar. Die Faktoren, welche die Gültigkeit von Angaben nachteilig beeinflussen, wirken interaktiv bzw. als bedingte Einflüsse je nach Kontext. Hierzu ein Beispiel: Es ist nicht einfach die Tatsache, entgegen den Erwartungen für “gute Staatsbürger” der Wahl ferngeblieben zu sein, die häufig die falsche Behauptung einer Wahlbeteiligung zur Folge hat; erst bei höherer Integration in die jeweilige Gemeinde bzw. Nachbarschaft stellt sich dieser Effekt, dann aber mit ziemlicher Stärke, wirklich ein. Ein weiteres Beispiel: Guter Rapport zwischen Interviewer und Befragtem erhöht die Chance unrichtiger Antworten, die aber der Erhaltung des guten Rapports förderlich sind. Ist jedoch die soziale Distanz zwischen den Partnern im Interview gering – Distanz für sich hatte ja wenig Wirkung –, dann wirkt sich guter Rapport erst recht nachteilig auf die Qualität solcher Angaben aus, die den Rapport als Verletzung “sozial wünschbarer Angaben” (social desirability) stören könnten (129). Als Fazit kann geschlossen werden, daß die in der Literatur behaupteten Störfaktoren gewöhnlich erst bei höherer Intensität der Ausprägung durchschlagen, daß aber ein Interpretationsspielraum bei Sachverhalten (“weiche” Daten) früher und stärker wirkt, als selbst unter Fachleuten angenommen wird.

Die wichtigste einzelne empirische Untersuchung der Gültigkeit von Angaben bei Interviews war bisher immer noch die “Denver Validity Study” des Jahres 1949 (130). Einer der damaligen Autoren, Don Cahalan, hat jetzt das damalige Material noch einmal ausgewertet – die Daten waren übrigens lange unter Umständen verschollen, die einem Bauernschwank nahe kommen – und der Bericht hierüber dürfte der wichtigste der hier betrachteten Aufsätze zum Thema Gültigkeit sein (131). In der damaligen Erhebung war die Angabe zu 15 Fragen an Hand schriftlicher Quellen überprüft worden. Wegen der Bedeutung dieser Erhebung seien die allgeminsten Befunde hier noch einmal wiederholt (132):

Frage	Angaben:			
	Richtig	Übertreibung	Untertreibung	Anderes
Besitzer eines Telephons	98%	1%	1%	—
Eigentümer der Wohnung	96	3	1	—
Automobilbesitzer	94	3	0	3
Alter	92	4	4	0
Führerscheibesitz	88	10	2	0
Leserkarte Volksbücherei	87	9	2	2
Teilnahme an Präsidentschaftswahlen	86	13	1	0
Eintragung ins Wahlregister, Lokalwahlen 1943-48	82	16	2	0
Präsidentschaftswahl 44	73	23	2	2
Bürgermeisterwahl 47	70	28	1	1
Vorwahlen (Primary) 48	69	21	3	7
Parlamentswahlen 46	69	19	2	10
Stadtratswahlen 47	60	28	2	10
Beitrag zu wohltätigen Vereinigungen 1948 (133)	56	34	0	10

Für die Analyse, die dem neuen Bericht zugrunde lag, wurden Angaben über Sachverhalte ausgewählt, die sich als relativ störanfällig erwiesen hatten: Angaben über die Teilnahme an den fünf erfaßten Wahlen, über die Leserkarte, den Besitz eines Führerscheins und den Beitrag zu wohltätigen Vereinigungen. Dies war angemessen, weil es in dem zweiten Bericht nicht mehr um die allgemeine Verlässlichkeit eines Merkmals ging, sondern um die Ableitung allgemeiner Sätze zur Gültigkeit von Angaben im Interview; störanfälligere Angaben sind wegen der größeren Variabilität der abhängigen Variablen geeigneter, Bestimmungsfaktoren zu identifizieren.

Die Korrelationen zwischen ungültigen Angaben und den üblichen demographischen Faktoren erweisen sich auch in dieser Untersuchung als durchweg unerheblich. Frauen geben etwas häufiger gültige Antworten, aber der Unterschied nach Geschlecht ist gering; und die nicht sehr starken Beziehungen zwischen Gültigkeit und Alter ergeben kein deutbares Muster. Ähnlich gering ist der Erklärungswert solcher im Folklore des Interviews für wichtig gehaltener Variablen wie Status des Befragten oder der Einfluß einzelner Interviewer. Auch *Don Cahalan* zeigt auf, daß diese Variablen erst in Kombination eine größere Aussagekraft erhalten, daß mithin die Effekte spezifischer als sonst formuliert sind. Die wichtigsten zusätzlichen Aussagen beziehen sich auf den Zusammenhang zwischen der Häufigkeit eines Phänomens und der Wahrscheinlichkeit ungültiger Antworten — ein Zusammenhang, auf den zuerst *Parry* und *Crossley* 1951 hinwiesen. *Cahalan* deutet solche Zusammenhänge mit der "Hypothese kognitiver Verzerrung": Befragte würden nicht nur dem Interviewer das Bild eines guten Mitbürgers vorspielen, sondern ihre eigenen Wahrnehmungen und Erinnerungen im Hinblick auf ihren Selbstrespekt verzerren. Daraus folgt als praktische Strategie die Notwen-

digkeit, Untersuchungsthemen so zu konzipieren, daß diese Sperre nicht aktiviert wird (134).

Mehrfache Befragungen der gleichen Personen – ob es sich um ein Panel i.e.S. handelt oder nicht – sind in modernen, sehr mobilen Sozialsystemen schwierig, und demgemäß hat sich in der Praxis ein ganzer Katalog von Vorgehensweisen und Notlösungen entwickelt. Diese sind durchweg intellektuell trivial, aber doch von größter Bedeutung für die Praxis. Zu den während des Beobachtungszeitraums erörterten Notlösungen gehören rollierende Samples (135) und Formen der Ersetzung von Fällen; zu den Vorgehensweisen schon beinahe detektivische Techniken. Nach einem Aufsatz, der in größerem Umfang unveröffentlichtes Material verwertet, konnte in 10 Untersuchungen – bei allerdings sehr, sehr hohem Einsatz von Mitteln – Ausschöpfungsraten der Stichproben erzielt werden, die höher als die üblicher Wahrscheinlichkeitsauswahlen mit bisher noch nicht befragten Fällen waren; dies relativiert bisherige, andersartige Erfahrungen (136). Allerdings wird hier wohl auch die Problematik der veränderten Bedeutung von Indikatoren unterschätzt (137). Dies ist leider bei der Analyse von Paneldaten und anderen Zeitvergleichen üblich: Die Analyseprobleme werden als statistische Fragen verstanden (138). Selbst dann sind die Beiträge zur Lösung der besonders schwierigen Analyseprobleme selten sehr informativ; *Goodman's* Pionierarbeiten haben keinen gleichwertige Nachfolge gefunden, und so gibt es noch viele offene Fragen.

Unter den in der Umfrageforschung seltenen quasi-experimentellen Techniken hat die "Technik der falschen Briefe" (Lost-Letter Technique) einiges Interesse gefunden, weil es sich um ein äußerst kostengünstiges und zeitsparendes Verfahren handelt, das auf verdeckte Weise eine erste rohe Schätzung der Stärke einer Einstellung zu Personen oder Institutionen erlaubt (139). An ausgewählten Stellen eines Gebiets werden bereits frankierte Briefe scheinbar verloren; die Anschriften dieser Briefe variieren, und aus der unterschiedlichen Bereitschaft der Finder zum Einwerfen der gefundenen Briefe wird auf die Beurteilung des Objektes in einer Bevölkerung geschlossen. Während des Beobachtungszeitraums kam es zu einer Kontroverse über die Aussagekraft eines Experiments, mit dem eine Wahlvorausage auf der Grundlage der "Technik der falschen Briefe" versucht wurde (140). Die Grundannahme des Erfinders der Technik, Stanley Milgram, daß die Bereitschaft zum Einwerfen der Briefe ein Ausdruck der Einstellung zu einem Objekt ist, konnte teilweise bestätigt werden – teilweise, weil eher die Ablehnung als die Bevorzugung eines Objektes gemessen werden kann (141). Seither ist die Technik weiter verfeinert worden, etwa durch Bestimmung der Kriterien für Orte, an denen Briefe "verloren" werden, aber sie bleibt doch ein recht grobes Maß, das sich vielleicht am ehesten zur Messung bei sehr umstrittenen Objekten eignet (142). Hierfür müßte man jedoch mehr Erfahrungswerte über Rücklaufquoten wissen – sie scheinen meist zwischen 25% und 35% zu liegen –, um die Ergebnisse einer bestimmten Untersuchung beurteilen zu können.

Experimentelle Versuchsanordnungen wurden in der Umfrageforschung vor allem zur Klärung methodischer Probleme der Fragebogen-Formulierung für mündliche Befragungen und zur Verbesserung der Ausschöpfungsquote bei postalischen Befragungen angewandt. Während des Beobachtungszeitraums erschienen zur Fragebogengestaltung lediglich Aufsätze, in denen Eigenschaften bereits bekannter Sonderformen von Fragen, wie des "Kartenspiels" (card sorting), untersucht wur-

den (143). Zu den wiederholt geprüften Alternativen zum mündlichen Interview gehört in den USA das Telephoninterview, dessen Aussagekraft allerdings wiederholt bezweifelt wurde. Diese sollte in einer Untersuchung geprüft werden, in der 50% der Befragten bereits ein Jahr zuvor in einer Wahrscheinlichkeits-Erhebung interviewt worden waren, während 50% in einer neuen Stichprobe bestimmt wurden (144). Das Ergebnis: "Die Gültigkeit der Telephoninterviews erscheint ebenso hoch wie die mündlicher Befragungen. Die Ausschöpfungsraten für Telephoninterviews sind ähnlich denen bei mündlichen Erhebungen" (145).

Die meisten experimentell angelegten Untersuchungen befassen sich mit den postalischen Befragungen. So wird geprüft, welche Art der Bezahlung des Befragten zu einer wesentlichen Verbesserung des Rücklaufs führen könnte (146). Inzwischen kann eine Summe der vielen empirischen Untersuchungen zu postalischen Befragungen gezogen werden, bei der sich einige der Handwerksregeln als nicht belegt oder sogar als widerlegt erweisen (147). So hat die Zusicherung strikter Anonymität keinen eindeutig positiven Effekt; gelegentlich wirkte ein persönliches Anschreiben – durch das ja Anonymität in Frage gestellt ist – günstig. Die Betonung der sozialen Nützlichkeit einer Erhebung bleibt wirkungslos und überraschend hat eine Verlängerung des Fragebogens nicht durchweg eine Verschlechterung der Antwortbereitschaft zur Folge. Positive Einflüsse hatten demgegenüber vorbereitende Kontakte mit den Befragten (etwa telephonische Ankündigungen der Befragung), die Art des Absenders (sponsorship effect), Zahlung einer kleinen Summe (etwa 1,50 DM), mehrfache und variierende Erinnerungsschreiben sowie die Art der Frankierung der Briefe. Auf diesem Gebiet der postalischen Befragung, wie auch bei Telephoninterviews, hat sich während der letzten 10 Jahren eine wesentliche Weiterentwicklung des handwerklichen Wissens ereignet.

Probleme der Datenanalyse sind für das Public Opinion Quarterly wesentlich weniger bedeutend als für die anderen untersuchten Zeitschriften. Dabei ist zudem noch ein besonderer Akzent auf praktischen Fragen der Auswertung auffällig. Hierzu gehört etwa die Frage nach den besten Techniken, um besonders früh ein Wahlergebnis voraussagen zu können (148), oder die Technik der Analyse nach Alterskohorten (149). Der für Praktiker vielleicht wichtigste Aufsatz ist eine Erklärung zweier Computer-Programme – Automatic Interaction Detection (AID) und Multiple Classification Analysis (MCA), von denen AID inzwischen ziemlich verbreitet ist –, mit denen automatisch diejenigen Variablen identifiziert werden können, welche den größten Effekt auf die abhängige Variable haben (technisch: zur stärksten Reduktion der Variabilität führen) (150). Werden diese und verwandte Techniken – wie Tree Analysis oder AUSPICK – gedankenlos angewandt, dann führen sie zwar zu Tabellen mit hohen Signifikanzgraden, haben aber einen gedankenlosen bloßen Induktionismus zur Folge; und diese Folge ist in der Tat eingetreten. Leider haben Aufsätze, aus denen an sich die Notwendigkeit zu großer Vorsicht bei der rein induktiven Manipulation großer Datenmengen abzuleiten wäre, insbesondere auf jüngere Sozialforscher keine Wirkung gehabt (151).

Methodische Probleme der Sekundäranalyse und deren Voraussetzungen, Datensammlungen, waren kein wichtiges Thema. Die bei weitem wichtigste Veröffentlichung in diesem Bereich war ein Buch (152). Leser der Zeitschrift erfuhren über diese Themen vornehmlich durch Berichte über Untersuchungen mit inhalt-

lichen Zielsetzungen (153) und wurden nur einmal über die neuen Auswertungsmöglichkeiten durch Datenbanken informiert (154). Wahrscheinlich folgt dies aus der Ausrichtung von POQ auf die Datensammler selbst.

Skalierung ist eines der konstanten aber Nebenthemen der meisten Zeitschriften für Sozialforschung, und auch des Public Opinion Quarterly. Meist werden hier einzelne Skalen erörtert – wie etwa eine empirische Untersuchung der Tendenz zur Überbetonung von Extremen bei zwei Typen von Skalen: “bipolar reversal scale” (einer Skala, deren Endpunkte durch einander entgegengesetzte Aussagen markiert werden) und “agreement scale” (praktisch wie ein Item einer Lickert-Skala) (155). Oder es wird untersucht, ob die “Roper Opinion Leadership Scale” anfällig gegenüber der Fehlerquelle “response set” ist – der Neigung von Befragten, unangesehen des Inhalts nach einem Schema (nur ‘ja’; nur ‘nein’; auf ein ‘ja’ folgt ein ‘nein’; etc) zu antworten (156). Neben der Auseinandersetzung mit Maßen, die als Standardformen in der Umfrageforschung benutzt werden, werden technische Fragen der Konstruktion von Skalen aufgegriffen, wie die Auswahl von Indikatoren. An empirischem Material wird aufgezeigt, daß mehr Sorgfalt als üblich darauf zu verwenden ist, vor der eigentlichen Item-Analyse die Bedeutung der vorläufigen Sammlung von Indikatoren zu prüfen, aus der dann die Aussagen für eine Item-Analyse ausgewählt werden (157).

Die in der Umfrageforschung bevorzugten Typen von Skalen werden nach den Verfahren von Lickert und von *Guttman* konstruiert; und da zu Lickert Skalen methodisch nicht mehr allzuviel zu sagen ist, werden von Methodologen vornehmlich die Probleme der *Guttman* Skalen erörtert. Gewöhnlich werden die Probleme dieses Verfahrens rein technisch behandelt, wie hier in einer Verfahrensweise zur Bestimmung derjenigen Ordnung von Kriterien (Items) einer Skala, die zu der geringsten Zahl von Fehlern gegenüber dem Modell führt (158). Gerade dies sei eine irrige Sicht der Problematik, argumentiert *John P. Robinson* vom Survey Research Center in Ann Arbor, da *Guttman* sein Verfahren sehr restriktiv verstanden habe als Modell für Sachverhalte, wo das Verhalten besonders strukturiert ist (159). Indem *Robinson* die Problematik des Gütekriteriums für die Eignung eines Materials, gemäß dem Modell von *Guttman* geordnet zu werden, näher untersucht (coefficient of reproducibility), kommt er zu dem Schluß: “Es sei behauptet, daß *Guttman*-Skalen am ehesten für solche Verhaltensweisen gefunden werden, die in einer Gesellschaft besonders stark strukturiert sind – wie soziale Distanz, Organisationshierarchien oder evolutionäre Stufen . . . Sie sind wahrscheinlich meist ungeeignet, um die Masseneinstellungen zu beschreiben, mit Ausnahme solcher Sachverhalte die in einer Umwelt besonders starr definiert sind (wie z.B. solche, denen eine Komponente sozialer Distanz immanent ist, wie Sexualverhalten oder der Gebrauch von Vulgaritäten)” (160). Diese Argumentation ist einleuchtend und stimmt in der Tat mit den frühen Darstellungen *Guttman*'s überein; die Implikationen dieser Konzeption von *Guttman* wurde dann im Verlauf der Popularisierung vollkommen undeutlich. Nach diesem Aufsatz von *Robinson* muß man zu dem Schluß kommen, daß die Beliebtheit der *Guttman*-Skalen in der Umfrageforschung auf einem Irrtum beruht und ihre Anwendung meist ungerechtfertigt ist.

Nach der Situation des Interview selbst, ist im Prozeß der Umfrageforschung das Zuweisen der Antworten zu Kategorien, also das Verschlüsseln oder Coden, die bedeutendste Schwachstelle dieses Erhebungsverfahrens. Angesichts dieser

Bedeutung des Verschlüsseln für die Qualität der Daten ist es zunächst verwunderlich, daß dieses Thema so selten aufgegriffen wird. Da wird zunächst untersucht, mit welchen Fehlern die Aufgabe des Interviewers, eine konkrete Antwort einer abstrakteren Kategorie zuzuweisen (wie dies bei vorformulierten Fragen notwendig ist), bei einem speziellen Komplex verbunden ist: Der sogenannten "Felderschlüsselung" von Angaben zu Fragen nach dem Beruf (161). Hier hat der Forscher als wesentliche Einwirkungsmöglichkeit nur die Gestaltung des Fragebogens selbst, und gute Antwortkategorien zu formulieren ist eine handwerkliche Fähigkeit, die sich erst über längere Zeit im Umgang mit Interviewern entwickeln läßt; sie geht deshalb vielen Akademikern ab. Für die in den Instituten selbst erfolgenden Verschlüsselungsarbeiten kann dagegen der Forscher die Fehler kontrollieren, und die Verfahren hierzu sind bekannt. In einer Untersuchung wird ein Verfahren entwickelt, mit dem die Kosten der Prüfung von Verschlüsselungsfehlern für "Instituts-Verschlüsselung" gesenkt werden sollen, da für die Anwendung optimaler Verfahren durchweg keine Mittel zur Verfügung stehen (162).

Die schwächste Stelle des ganzen Verschlüsselungsvorgangs ist mit Abstand die Verschlüsselung offener Fragen. Die in der Praxis der Umfrageforschung entwickelten Codes taugen meist nicht sehr viel, und die Zuordnung der wörtlichen Niederschriften ist mit höheren Fehlern behaftet. Gelänge die Ausbildung eines automatischen Verfahrens, mit dem Codes entwickelt würden und dann auch eine automatische Zuordnung der wörtlichen Angaben erfolgte, dann wäre dies eine beachtliche Verbesserung der Aussagekraft von Umfragen. Am Material einer Befragung über "Glück", also einem recht anspruchsvollen Gegenstand, soll die Möglichkeit demonstriert werden, ein automatisiertes Verfahren der Inhaltsanalyse für die Verschlüsselung von Umfragedaten zu nutzen (163). Der Versuch führt zu keinem eindeutigen Ergebnis, und aus eher zufälligen Umständen wurden die Arbeiten seither nicht weitergeführt. So blieb es bisher beim alten: Die Verschlüsselung offener Fragen ist mit Fehlern und hohen Verlusten an Informationen verbunden – ein Umstand, der in Analysen oft nicht berücksichtigt wird.

5.

Diese Übersicht war als Auswertung der Zeitschriften als dem zentralem Forum für methodologische Fragen konzipiert. Für einige wichtige Entwicklungen erwiesen sich jedoch zumindest diese Zeitschriften als unvollständig. Zu diesen Themen, die überraschend auch im englischen Sprachbereich kaum Autoren für Aufsätze fanden, gehören die methodischen Probleme der Sekundäranalyse (vgl. Anmerkung 152 als Korrektur). Nur als Korrektur, um die Nützlichkeit dieser Arbeit als Zugang zur aktuellen Erörterung von Methodenfragen zu erhalten, sei auf einige weitere Themen verwiesen; sie können allerdings nicht so systematisch wie bei der Auswertung der Zeitschriften berücksichtigt werden.

Da ist zuerst die zunehmende Bedeutung sogenannter Aggregatdaten und ihrer Auswertung zu beachten. "Aggregatdaten" und/oder ökologische Daten sind die nicht sehr zweckmäßigen Bezeichnungen für Daten, die als Nebenprodukt des Verwaltens aller Art anfallen (164). Entsprechend wäre die Bezeichnung "Prozeß-

daten" (process produced data) vorzuziehen. Im 19. Jahrhundert bis einschließlich den 20er Jahren waren es diese Daten, mit denen die empirischen Sozialforscher bevorzugt arbeiteten. Nachdem seither die Sozialforscher zu ihren eigenen Datensammlern wurden, erlangten in den sechziger Jahren Prozeßdaten eine zunehmende Bedeutung vor allem in der politischen Soziologie und der Stadtsoziologie. Dazu trug die Mehrebenenanalyse als ein neuer Forschungsansatz bei, jedoch auch die neue Zugänglichkeit interessanter Daten. An einer wirklich ausgebildeten Methodenlehre für diese Daten fehlt es noch, aber in einigen Anthologien ist doch eine teilweise Summe der Erfahrungen mit dieser Kombination von Kenntnissen in Datenaufbereitung, Nutzung von Analysepaketen für Computer der dritten Generation und deskriptiver Statistik dargestellt (165).

Die Nutzung von Prozeßdaten für die Sozialforschung steht im Zusammenhang mit der Entwicklung "sozialer Indikatoren". Diese letztere Entwicklung ist jedoch primär nicht aus den Sozialwissenschaften heraus entstanden. Wesentlich als Folge der wachsenden Staatstätigkeit und der Zunahme bürokratischer Veranstaltungen, wuchs die Datenproduktion zu einer Abbildung des Alltags. Aus welchen Gründen immer verläuft parallel hierzu eine Ausdehnung der Ansprüche an Regierungen und Verwaltungen, ordnend in den Alltag einzugreifen bis hin zu der Erwartung, daß politische und administrative Institutionen für "Lebensqualität" verantwortlich sein sollen. "Soziale Indikatoren" sind nun die Daten – überwiegend prozeßproduzierter Art –, welche politische und administrative Instanzen in die Lage versetzen sollen, den Zustand des Gemeinwesens im Hinblick auf Gestaltungswünsche zu beurteilen. Eine kürzere Definition lautet: Soziale Indikatoren sind Daten mit normativen Implikationen. Demgemäß ändert sich die Vorstellung darüber, was ein sozialer Indikator ist, aber inzwischen haben sich Gruppen von Sozialwissenschaftlern der Aufgabe angenommen, Systematik in die Formulierung der Kataloge sozialer Indikatoren und ihrer Deutung hinein zu bringen (166). Mit der Ausbildung kompletter Informationssysteme ist zu erwarten, daß über die sozialen Indikatoren i.e.S. hinaus prozeßproduzierte Daten für die Forschung verfügbar werden.

In den sechziger Jahren nahm die Zahl von Berichten über internationale Vergleiche rasch zu. Die Hauptmotive waren, ein tertium comperationis für die Daten des eigenen Landes zu gewinnen und die Variationsbreite der Beobachtungen zu vergrößern. Dabei stellten sich zahlreiche forschungstechnische Fragen, die sich aber im wesentlichen nur als Fortführung bereits vertrauter Probleme insbesondere der Datensammlungen erwiesen (167). Zusätzlich wurden einige Probleme deutlich, die sich durchaus auch für Sozialforschung im eigenen Lande ergeben, dort aber nicht so deutlich werden. Ein Beispiel ist der Nachweis aufgrund der Daten einer international vergleichenden Umfrage, daß die in der Fachliteratur öfters untersuchte Antwort "Weiß nicht" (don't know) je nach Land anders zu interpretieren ist, da sie auch eine "response set" Komponente einschließt (168). Während in den USA eher geraten als auf Meinungslosigkeit ausgewichen wird, gilt für französische Befragte tendenziell das Gegenteil. Die schwierigeren und wichtigeren Probleme ergeben sich bei der Forschungsplanung und insbesondere der Deutung der Ergebnisse. Die früher (und zum Teil heute noch) verbreitete Neigung, Unterschiede zwischen den Daten auf verschiedenen Ländern zu deuten als Ausdruck der "Kultur" oder sonstigen Eigenheit eines Landes, führte zu feh-

lerhaften Zurechnungen von Ursachen und Wirkung; diese Zurechnung ist bei international vergleichenden Erhebungen eher noch schwieriger (169).

Inzwischen stagniert die international vergleichende Sozialforschung als ein eigener Forschungsansatz. Die geringeren Mittel für empirische Forschung wirkten sich auf eine so besonders teure Art der Datensammlung besonders stark aus. In der Zwischenzeit wurden jedoch umfangreiche Datensammlungen für Sozialforscher auch zu Ressourcen für international vergleichende Sekundäranalysen, und diese Analysen sind sehr viel kostengünstiger als Primärerhebungen. Zudem wurden in den Jahren der großen internationalen Primärerhebungen Netzwerke persönlicher Bekanntschaften aufgebaut, die eine direkte Abstimmung der Erhebungsarbeiten in verschiedenen Ländern erlauben, ohne sich auf zentral geplante Untersuchungen mit hohen Koordinationskosten verlassen zu müssen. Die Stagnation hat eher intellektuelle Gründe: Nach Jahrzehnten international vergleichender Individualzählungen muß zugegeben werden, daß die Unterschiede speziell zwischen verschiedenen Industriegesellschaften auf dieser Ebene der Beobachtungen nicht ausreichend über die Unterschiede zwischen den Ländern insgesamt und ihre Veränderungen aussagen.

Zu dieser Beobachtung, daß die Unterschiede in den Daten für verschiedene Länder insbesondere für Untersuchungen mit makro-soziologischen Problemstellungen nicht die erhoffte Aussagekraft haben, kommt eine weitere Beobachtung. Inzwischen liegen für verschiedene Industriegesellschaften Daten über einen Zeitraum von über 20 Jahren vor, und deren Veränderung kann zum Teil nur so gedeutet werden, daß ein Teil der Unterschiede zwischen den Ländern eher Stimmungen als die Eigenschaft einer Bevölkerung ausdrückt. So galt es als ausgemacht, daß ein Hauptunterschied zwischen den USA und Frankreich die Bereitschaft der Bevölkerung in den Vereinigten Staaten sei, den politischen Institutionen ihres Landes einen hohen Vertrauensvorschuß zu gewähren, ungeachtet der Beurteilung konkreter Amtsinhaber und Parteien. Aus dieser optimistischen Grundhaltung gegenüber dem eigenen politischen System in den USA wurde die Funktionsfähigkeit der Institutionen in den USA unmittelbar abgeleitet. Inzwischen erweist sich jedoch, daß in den Ländern zum jeweiligen Zeitpunkt zu einem erheblichen Grad Stimmungen und nicht bleibende Eigenschaften gemessen wurden (170). Diese Beobachtung hat selbstverständlich Bedeutung über die international vergleichende Forschung hinaus für Individualzählungen überhaupt, insbesondere für Befunde der Meinungsforschung.

6.

Gibt es eine Krise der quantitativen Sozialforschung? Sicherlich nicht als institutionalisierte Tätigkeit der Sozialwissenschaftler: Empirische Forschung ist auch in der Bundesrepublik zu einer Selbstverständlichkeit geworden, weitgehend ungeachtet wissenschaftlicher und politisch-ideologischer Unterschiede zwischen den Soziologen. Das "Informationszentrum für sozialwissenschaftliche Forschung" führt seit Ende der sechziger Jahre Erhebungen über Forschungstätigkeit durch und berichtet ein fortlaufendes Ansteigen der Zahl empirischer Projekte (171). Meist handelt es sich allerdings um kleine Projekte, die entgegen verbreiteten Vorstellungen über die mangelnde Praxisorientierung der Soziologie stark anwendungs-

bezogen sind. Eine interdisziplinäre Orientierung herrscht bereits vor, wobei eine auch soziologische Orientierung häufiger ist als die Beteiligung der Soziologen selbst – womit sich dann die Soziologie in der Forschungspraxis als erfolgreicher Exportartikel ausweist. Auch die Feuilleton-Kritik an der Art der empirischen Forschung ist ohne Einfluß geblieben: Nach wie vor ist das Interview die bevorzugte Technik der Datensammlung und die Tabelle die übliche Art der Darstellung (172).

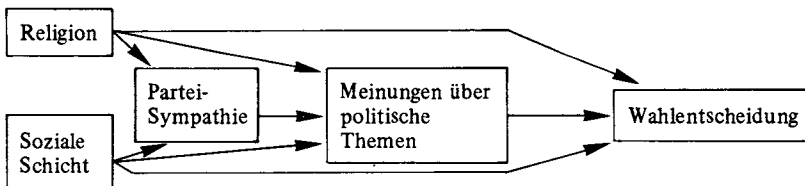
Dieser Zustand der Sozialforschung auch in der Bundesrepublik sagt allerdings in erster Linie nur etwas aus über den inzwischen erreichten hohen Institutionalisierungsgrad der Sozialforschung. Bei dieser Ausdehnung der Sozialforschung weit über die Soziologie hinaus, der Konzentration auf Anwendung, meist auf "soziale Probleme", dem Vorherrschen von Berufsanfängern in der praktischen Forschung und der routinehaften Ausbildung in Forschung als Erlernen von Buchwissen kann nicht verwundern, daß diese Sozialforschung durchweg Routinecharakter hat. Das ist an sich noch kein Grund für Kritik: Gerade daran, welche Themen von Forschungsfragen zu Selbstverständlichkeiten werden, läßt sich die fortschreitende Institutionalisierung, und im innerwissenschaftlichen Sinne der Fortschritt einer Disziplin, ablesen. Wenn demgegenüber in der Öffentlichkeit gelegentlich von Sozialwissenschaftlern die Auffassung vertreten wird, das Fehlen größerer Neuerungen und heftiger Grundsatzdebatten sei ein Indiz für die Stagnation einer Disziplin, dann wird ein Kriterium des Kulturbetriebes mit Wissenschaft verwechselt. Die Kritik muß an einer anderen Eigenschaft der institutionalisierten Sozialforschung in Europa ansetzen: Dem unzureichenden handwerklichen Können insbesondere bei der Datensammlung. Demgegenüber hat die Verbreitung von Computern und Analysepaketen die erwartete Wirkung gehabt, die Analyse um die anspruchsvolleren statistischen Techniken zu erweitern.

Gerade hier entwickelt sich eine wichtige methodologische Erörterung. An sich hat sich die Methodenliteratur, so wie sie insbesondere in den zunächst untersuchten sieben Zeitschriften bzw. Jahrbüchern repräsentiert ist, gegenüber der Forschung als alltäglichem Geschehen weitgehend verselbständigt. Die für Forschung als Handwerk bedeutsamen Erfahrungen und Untersuchungen haben lediglich im "Public Opinion Quarterly" ein regelmäßiges Forum. Sonst konzentriert sich die methodologische Erörterung in einem überraschenden Maße auf Techniken der Datenmanipulation. Speziell durch die Analysepakete wie SPSS, OSIRIS, CROSSTABS oder DATA-TEXT wird ein Teil dieser Verfahren der Datenmanipulation zurückvermittelt in die Forschung. Aus der Forschung selbst scheinen aber nur in geringem Umfang die Anregungen zur Weiterentwicklung der Datenanalyse zu kommen. Methodologie in ihrer heutigen Erscheinungsform insbesondere in den USA ist zu einem Ensemble von Spezialitäten geworden. Und diese in Spezialitäten zerfallende Spezialität wird mitbestimmt durch das Bedürfnis nach Novitäten. Insbesondere bei der Entwicklung neuer Assoziationsmaße, von Skalierungsverfahren und von Modellen scheinen die Autoren häufig mehr einander beeindruckt zu wollen, als die Sozialforschung zu verbessern (173).

Es gibt ausreichend Probleme der praktischen Forschung als Information über Realität, welche die Erfindungsgabe der so hoch spezialisierten Methodologen anregen könnten. Durch Datenbanken und den Aufbau von Berichterstattungen über "soziale Indikatoren", und durch deren Weiterentwicklung zu Komponenten von

Informationssystemen, ergeben sich neue Probleme der Datenmanipulation. In diesem Zusammenhang stellen sich interessante Fragen bei der Mehrebenen-Analyse. Bei der Analyse der vorherrschend benutzten Individualzählungen sind dies aus zwei Umständen in den Sozialwissenschaften ergebenden Probleme ungenügend reflektiert: die Probleme der Multikollinearität und der Überdeterminiertheit von Verhalten. Multikollinearität als Problem ist eben durch Techniken der Partialkorrelationen nur vordergründig berücksichtigt, und Überdeterminiertheit wird durch Techniken wie "tree analysis" als bloße Irritation erfaßt.

Die Soziologie als eine empirische Disziplin ist heute noch mehr als früher dem Modell der vorgeblich erfolgreichsten Naturwissenschaft verpflichtet: Der Physik. Vor der Überflutung der Physik mit Daten wurde dort generell unterstellt, daß letztlich einige wenige "letzte" Naturkräfte (üblicherweise werden vier genannt) die zu erklärenden Erscheinungen bestimmten. Dem entsprach eine Forschung, die Wirkkräfte als einzelne Variable zu isolieren versucht. Die Tabelle der Sozialforscher, wenn deren Charakter in einer Maßzahl insgesamt ausgedrückt werden soll, ist eine ex post Isolierung von Faktoren. Nun erweist es sich jedoch, daß die empirische Forschung dieser Art über Erklärungen der Variation bei den abhängigen Variablen von 30% der Variation nur selten hinauskommt. Sehr viele der zu erklärenden Erscheinungen erweisen sich eben als mehrfach determiniert, und die Interaktion zwischen Faktoren hat häufig den Charakter bedingter Kausalität. Hinzu kommen die Erscheinungen, die theoretisch mit den Begriffen funktionale Äquivalenz und Multifunktionalität ausgedrückt werden: Daß also Konkrete je nach Bedingungs Zusammenhang anders wirken können, oder daß sich die gleiche Wirkkraft durch unterschiedliche Konkrete ausdrückt. Biologische Phänomene kommen dem hier skizzierten Zustand nahe, der hier am Beispiel der Erklärung des Wahlverhaltens schematisch vereinfacht dargestellt sein soll (174):



Die Pfadanalyse ist ein Versuch, diesem interaktiven Charakter bei der Verursachung von Erscheinungen Rechnung zu tragen. Es gibt jedoch Anlaß zur Skepsis, wenn inzwischen Erfahrungen vorliegen, daß Pfadanalysen mit mehr als fünf Faktoren selten gelingen, die methodologische Literatur dies aber nur als ein mathematisch-statistisches Problem reflektiert. Angemessener wäre die Entwicklung weiterer Verfahren und Modelle, die dem oben skizzierten Charakter vieler sozialwissenschaftlicher Erklärungsgegenstände gerechter würden.

Dies sind selbstverständlich Probleme, die sich die Sozialforschung erst einhandelte, nachdem sie als Sozialbeschreibung erfolgreich war und es ihr ferner gelang, mit nicht sehr anspruchsvollen Verfahren viele Erscheinungen des Alltags auf die Variation einiger Bestimmungsgründe teilweise zu reduzieren. Seither erweist es sich jedoch als schwierig, über solche Erklärungen hinaus zu gelangen. Dies ist ein Aspekt dessen, was zu Beginn dieses Abschnitts als "Krise der quantitativen Sozial-

forschung" bezeichnet wurde. Inzwischen mehren sich im Zusammenhang mit der Diskussion über Meßniveaus, die bisher allerdings vornehmlich den Charakter statistischer Spezialfragen hat, die Anzeichen dafür, daß selbst in den bisher erfolgreichen Anwendungen ein Teil der produzierten Befunde Artefakte des Vorgehens gewesen sein könnten.

Ein Beispiel ist eine Re-Analyse der Untersuchungen über Wohnzimmer-Stile als Indiz für soziale Schichtung. Ende der sechziger Jahre wiederholten *Howard Schuman* und *Edward O. Lauman* Analysen, wie sie zuerst von *Sewell* in den dreißiger Jahren durchgeführt wurden. Hiernach sind auch heute noch Wohnzimmer-Stile Indizien für soziale Gruppen, im vorliegenden Fall unterschiedlicher Einstellungen zur Modernität. *Marcus Felson* kommt nun bei einer Neuanalyse der gleichen Daten zu dem Schluß, daß die signifikanten Befunde ein Artefakt statistischer Manipulationen sind, der Transformation einer Variablen mit drei Antwortkategorien in drei 'dummy'-Variable (175). Über diesen Fall hinaus generalisiert *Felson*, daß 5% aller in Publikationen als signifikant bezeichneten Korrelationen in Wirklichkeit Artefakte der Vorgehensweisen seien (176).

"Alchemie in den Verhaltenswissenschaften" heißt eine Untersuchung, in der für eine größere Vorsicht bei der Anwendung solche aussagekräftiger statistischer Manipulationen wie AID (als Computer-Programm formalisiert), multiple Regression, Faktorenanalyse und mehrdimensionale Skalierung nach dem Verfahren nach *Kruskal* plädiert wird (177). Angewandt auf empirische Daten zeigt sich, daß bei einem nach Zufall verschieden unterteilten Datenmaterial das Verfahren AID zur Ermittlung besonders erklärungssträchtiger Variablen sehr instabile Ergebnisse erbrachte. Die Befunde multipler Regression waren dann besonders instabil, wenn die Zahl der unabhängigen Variablen relativ zur Zahl der Beobachtungen groß ist. In einer älteren Untersuchung hatten *McCarty* und *Fitzpatrick* mittels multipler Regression einen Test konstruiert, für dessen Indikatoren sie schließlich eine stolze Korrelation von .92 errechneten; bei Prüfung der Gültigkeit an einem zweiten Sample ergab sich phantastischerweise eine Korrelation von -.21! (178). Nach einer Prüfung der Berechnungen von Faktorenanalysen wurde ermittelt, daß in 46% weder eine Prüfung nach Gültigkeit noch Verlässlichkeit vorgenommen wurde, und daß Sampling-Variationen durchweg nicht berücksichtigt wurden (179). *Einhorn* warnt: ". . . in dem Maße wie Methoden und Techniken komplizierter werden, wird die Rolle der Theorie (bzw. inhaltlicher Aussagen) im Forschungsprozeß immer stärker zugunsten rein induktiven Vorgehens ignoriert (180). Und die methodischen Kautelen sowie die Unterstellungen in den Algorithmen sind sehr vielen derjenigen unbekannt, die Standardpakete für Analyse mit Hilfe moderner Computer benutzen.

Eine Revolution der Methodologie erwartete *McGinnis*, und diese hat sich sicherlich nicht ereignet. Rapide verändert hat sich dagegen die praktische Forschung, indem durch moderne Analysepakete und aufbereitete Datensätze auch Berufsanfänger Datenmanipulationen durchführen können, die früher nur erfahrenen Forschern vorbehalten waren. Zugleich sinkt zum Teil die Verlässlichkeit und Gültigkeit der Daten, was sich nicht zuletzt in der Konzentration von Aufsätzen zu diesem Thema im "Public Opinion Quarterly" ausdrückt. Damit eröffnet sich eine Schere zwischen verfügbaren Verfahren der Datenmanipulation und der Aussagekraft der Datenbasis. Wahrscheinlich ist dieser Zustand in den USA ausgeprägter

(181), aber auch in Westeuropa läßt sich beobachten, daß die Fähigkeit zur Nutzung hochentwickelter Verfahren der Datenmanipulation zunimmt, das Verständnis für die Aussagekraft der Daten jedoch bestenfalls stagniert.

Die methodologische Literatur hat andere Problemstellungen und ist durch einen nur mangelhaften Bezug zu den tatsächlichen Problemen der Forschung als Auskunft über Realität gekennzeichnet. Hinzu kommt bei jüngeren Sozialforschern eine zunehmende Hilflosigkeit gegenüber Daten, da der Umgang mit wirklichen Daten zugunsten von Buchwissen in der Ausbildung zu kurz kommt. Darüber hinaus sind in Taschenbuchveröffentlichungen – nicht nur in diesen, aber dort massenhaft – zahlreiche Fälle von extremer Voreingenommenheit im Umgang mit den eigenen oder anderen Daten zu beobachten (182). Dies ist dann nur noch äußerlich Sozialforschung, aber in Wirklichkeit keine Befragung der Realität mehr, sondern Illustration von Voreingenommenheiten. Daß hierfür die Äußerlichkeiten der Forschung als notwendig angesehen werden, ist dann allerdings indirekt ein Tribut an das Ansehen der Sozialforschung als Routine. Die letztere Entwicklung gibt Grund zur Sorge, da in Europa das methodologische Publikum als Korrektiv für solche elementaren Fehler noch nicht ausreichend vorhanden ist.

Die Methodologie der Sozialforschung befindet sich in einer Übergangssituation. Eine Krise im üblichen Sinne dieses Wortes ist es nicht – und wäre sie es, dann immerhin auf einem Niveau der Probleme, das viele andere Disziplinen erst einmal erreichen müssen. Krisenhaft könnte die Situation genannt werden, weil sich nicht nur manche als gesichert behandelten Befunde als Artefakte der Forschungstechniken erweisen werden, sondern auch Handwerksregeln sowie Manipulationstechniken als revisionsbedürftig herausstellen dürften. Die Problematik wird jetzt allerdings in der methodologischen Literatur zunehmend eingegrenzt. Es muß abgewartet werden, wann sie zum zentralen Thema wird. Selbstverständlich wird es auch dann weiterhin Methodologie als Spezialität geben, die gegenüber der Forschung als Klärung von Sachverhalten verselbständigt Neuheiten produziert.

Anmerkungen

- 1 Hierzu zwei Beispiele: 1. Zum Zeitpunkt des Erscheinens wurde die etwas hastig zusammengestellte Aufsatzsammlung *Theodor W. Adorno* u. Mitarb.: *Der Positivismusstreit in der deutschen Soziologie*, Luchterhand, Neuwied 1969, von der Publizistik als repräsentativ für die Zerklüftung der Soziologie in der Bundesrepublik behandelt; Studenten zogen es nicht selten vor, ihren Eindruck von Soziologie aus diesem Bändchen zu gewinnen statt aus Lehrbüchern oder Vorlesungen. Wen interessieren heute noch solche Themenstellungen wie die von *Habermas* "Gegen einen positivistisch halbierten Rationalismus"? , und wen haben sie eigentlich damals interessiert, falls er nicht von Studenten zur Stellungnahme gedrängt wurde? Welchen Einfluß haben die Ausführungen von Adorno über "Soziologie und empirische Forschung" je auf diese Forschung selbst gehabt? 2. Auch das Bändchen von *H.J. Krsymanski* und *Peter Marwedel* (Hrsg.): *Die Krise in der Soziologie*, Pahl Rugenstein, Köln 1975 sagt nichts zur soziologischen Forschung als institutionalisierte Tätigkeit; vielmehr wird das Krisenempfinden von Personen am Rande der Soziologie als Profession ausgedrückt. Dies alles war und ist eine Diskussion über die Soziologie und nicht der Soziologie.
- 2 An sich hat "Positivismus" in der Philosophiegeschichte bekanntlich eine feste Bedeutung, und bezeichnet eine dem "neo-Positivismus" diametral entgegengesetzte Position. Die Auffassungen von *Auguste Comte* über die Stellung der Soziologie unter den Wissenschaften

und über den Gang der Entwicklung sind dabei den Auffassungen Adornos nicht immer entgegengesetzt.

“Positivismus” in der Bedeutung als pejorative Bezeichnung für eine Soziologie als Erfahrungswissenschaft ist abgeleitet von einem privaten Sprachgebrauch in Frankfurt. Damit soll abgewertet werden, daß Tatsachen als Tatsachen behandelt werden, statt einer Verpflichtung der Soziologen darauf, über das Wünschenswerte zu spekulieren und aus diesen Spekulationen Imperative mit dem Absender Soziologie abzuleiten. Diese spezifischere Bedeutung hat sich dann abgeschliffen, als “Positivisten” in den Sprachschatz bundesdeutscher Feuilletonisten aufgenommen wurde. Übriggeblieben ist die Konnotation, daß “Positivisten” in einem naiven Sinne Empiristen seien. Das eben ist jedoch eine an der analytischen Philosophie orientierte Sozialforschung nicht. Somit bleibt auch in der verdünnten Bedeutung die Bezeichnung “Positivist” meist eine massive Irreführung.

- 3 Vgl. hierzu die Stellungnahmen von René König zur sogenannten Frankfurter Schule in René König (Hrsg.): *Handbuch der empirischen Sozialforschung*, 3. Aufl., Band 1, Enke, Stuttgart 1973, Einleitung; und Emile Durkheim: *Der Selbstmord*, Luchterhand, Neuwied 1973, Nachwort von R. König, insbesondere S. 473-500. König zeigt dabei auch die völlige Fehlinterpretation Adornos über die Bedeutung von Befunden für die Arbeitsweise von Durkheim auf. Diese Fehlinterpretationen und die Propagierung des unsinnigen Terminus “Positivismus” für eine Soziologie als Erfahrungswissenschaft ist jedoch einfach zu verstehen. Adorno kannte als Empirie wesentlich nur Umfragen als Auftragsforschung – so auch in seinem Institut – und behandelte diese praktische Nutzung einer Forschungstechnik als empirische Forschung überhaupt.
- 4 Dies drückt sich inzwischen in den Lehrbüchern von Autoren mit den unterschiedlichsten weltanschaulichen Orientierungen aus. Lehrbücher der empirischen Sozialforschung gehören zu den sozialwissenschaftlichen Schriften mit den höchsten Auflagen: Das Handbuch von René König erscheint trotz eines insgesamt sehr hohen Preises bereits in dritter Auflage. Peter Atteslander: *Methoden der empirischen Sozialforschung*, de Gruyter, Berlin 1975 erreichte in wenigen Jahren ebenfalls drei Auflagen, Renate Mayntz u. Mitarb.: *Einführung in die Methoden der empirischen Soziologie*, Westdeutscher Verlag, Opladen, erscheint jetzt ebenfalls in dritter Auflage. Ungewöhnlich rasch steigen auch die Verkaufszahlen der neuesten methodologischen Reihe, der “Studienskripten zur Soziologie”, Stuttgart mit inzwischen einem Dutzend Bänden. Selbst bei Schriften aus der “Frankfurter Schule”, wie dem Lehrbuch der Statistik von Ritsert und Becker sind nicht die Techniken als solche problematisiert, sondern die Einstellungen, mit denen sie benutzt werden. Die Sozialforschung als Technologie wird so etwas wie der allen Richtungen der Soziologie gemeinsame weiße Kittel.
- 5 Diese an sich offensichtliche Entwicklung wurde verdeckt durch die Behauptung der Trennung der Soziologen in Lager mit unterschiedlichen “Erkenntnisinteressen”. Dieses Wort verselbständigte sich rasch gegenüber der doch etwas eingeschränkteren Bedeutung bei Habermas; vgl. seine Aufsatzsammlung *Erkenntnis und Interesse*, Suhrkamp, Frankfurt 1968. Als ob es bei der Forschung als routinisiertem Verhalten für die Ergebnisse und Wirkungen primär auf das “Erkenntnisinteresse” des Forschers ankäme! Letztlich handelt es sich hier um den völlig untauglichen Versuch, die sozialen Wirkungen von Sozialforschung durch Kontrolle über die Motive bzw. Gesinnung der Sozialforscher selbst in den Griff zu bekommen. Die meist von Außen her vorgenommene Trennung in Lager mit angeblich unterschiedlichen Erkenntnisinteressen verdeckte die tatsächliche Gemeinsamkeit der Technologien und deren Nutzung.
- 6 Vgl. hierzu René König (Hrsg.): *Handbuch der empirischen Sozialforschung*, 2. Aufl., Band 2, Enke, Stuttgart 1969, Nachwort S. 1278.
- 7 Vgl. hierzu die frühen Schriften *Empirische Sozialforschung*, Institut zur Förderung öffentlicher Angelegenheiten, Frankfurt 1952 (über das Treffen in Weinheim, 14.-16. Dezember 1951); René König u.a. (Hrsg.): *Das Interview*, Verlag für Wirtschaft und Politik, Köln, 1. Auflage 1952; DIVO: *Praktikum der Meinungsforschung*, Frankfurt o.J. (1952). Der Eindruck wurde noch dadurch bestärkt, daß die damals jungen Sozialforscher, die in der Folgezeit die Empirie in der Bundesrepublik trugen, ihre Ausbildung in empirischen Verfahren in den USA, meist als Fulbright-Stipendiaten, erhielten.
- 8 Fragebögen wurden bereits für die Erhebungen der englischen Royal Commissions benutzt; Karl Marx entwarf bekanntlich im Londoner Exil für die deutschen Sozialdemokraten

einen standardisierten Bogen zur Erklärung des Wählerverhaltens; *Max Weber* war der spiritus rector der Erhebungen im Auftrag des Vereins für Socialpolitik über Anpassungsprobleme der Arbeiterschaft an die industrielle Produktion; usw. Siehe auch den Beitrag von *Heinz Maus* "Zur Vorgeschichte der empirischen Sozialforschung" in *René König*, *op. cit.*, 3. Aufl.; und den Aufsatz von *Paul Lazarsfeld* in *Tendances principales de la recherche dans les sciences sociales et humaines, Partie I: Sciences sociales*, Mouton, den Haag 1970.

- 9 Die Entwicklung der Beobachtungsverfahren erfolgte primär außerhalb der Sozialforschung i.e.S., insbesondere in der Volkskunde (zu ihrer Geschichte vgl. *Hermann Bausinger: Volkskunde*, DWG, Berlin, o.J., 1973), der Ethnographie und der frühen Marktforschung. Für ein Beispiel aus der Sozialforschung der Zeit vor dem 1. Weltkrieg siehe *Charles Booth: Life and Labour of the People of London*, 17 Bände, MacMillan, London 1892-1897. Der Pittsburgh-Survey und der Springfield Survey in den USA waren durch dieses Beispiel aus Europa inspiriert. Sie auch *René König* u. Mitarb. (Hrsg.): *Beobachtung und Experiment*, Verlag für Politik und Wirtschaft, Köln 1956.
- 10 Die erste bekanntgewordene Anwendung von Wahrscheinlichkeitsauswahlen auf menschliche Bevölkerungen wird *Arthur Bowley* bei einer Untersuchung von Armut 1908 zugeschrieben (Vgl. *Arthur L. Bowley* u. Mitarb.: *Livelihood and Poverty*, King, London 1915). Auch in Norwegen und Dänemark wurde mit Annäherungsformen an volle Wahrscheinlichkeitsauswahlen experimentiert. In den 20er Jahren ging dann eine frühere Assistentin von *Bowley* zu einer Vortragsreise in die USA. Dort wurden die Wahrscheinlichkeitsauswahlen zuerst von einigen Agricultural Experiment-Stations übernommen und fanden – nicht verwunderlich angesichts der engen Bindung dieser Forschungseinrichtungen an die Behörden in Washington – Beachtung durch Bundesbeamte in Washington. Es erfolgten wesentliche Weiterentwicklungen als Anpassung an die Verhältnisse an ein Land ohne amtliche Meldepflicht und mit sehr mobiler Bevölkerung. Diese Entwicklungen vornehmlich in Zusammenhang mit dem statistischen Bundesamt der USA wurden dann nach dem zweiten Weltkrieg nach Europa exportiert. Siehe zu dieser Entwicklung *Erwin K. Scheuch: Die Anwendung von Auswahlverfahren bei Repräsentativ-Befragungen*, Dissertation Köln, 1956. Der eigentlich amerikanische Beitrag zu den Auswahlverfahren ist das Quotenverfahren.
- 11 Vorläufer der quantitativen Inhaltsanalyse finden sich in der positivistischen Linguistik des 19. Jahrhunderts. Die systematische Auswertung von Tagebüchern etwa in der Psychologenschule um das Ehepaar *Bühler* darf ebenfalls als Inhaltsanalyse bezeichnet werden.
- 12 Die methodischen Grundlagen der Tests wurden in der Psychophysik entwickelt. Das erste Labor für diesen neuen Wissenszweig wurde von *W. Wundt* in Leipzig gegründet. Die Grundlegung für die behavioristische Psychologie erfolgte durch *Pawlow*. Zur Geschichte des quantifizierten Erfassens von psychischen Eigenschaften siehe die entsprechenden historischen Abschnitte in *H. Gulliksen: Theory of Mental Tests*, New York 1950; *J.P. Guilford: Psychometric Methods*, 2. Aufl., New York 1954; *A.L. Edwards: Techniques of Attitude Scale Construction*, New York 1957; *W.S. Torgerson: Theory and Methods of Scaling*, New York 1958; *P.F. Lazarsfeld: "Notes on the History of Quantification in Sociology"*, in *Harry Woolf* (Hrsg.): *Quantification – A History of the Meaning of Measurement in the Natural and Social Sciences*, Indianapolis 1961.
- 13 Die Soziometrie wurde nach dem 1. Weltkrieg in Österreich von *J.L. Moreno* entwickelt und nach dessen Übersiedlung in die USA dort im wesentlichen nur noch mathematisiert und dabei allerdings des bei *Moreno* wuchernden metaphysischen Beiwerks entkleidet. Siehe hierzu *Iri Nehnevajsa: "Sociometry – Decades of Growth"*, in *J.L. Moreno* (Hrsg.): *The Sociometry Reader*, Free Press, Glencoe Ill, 1960.
- 14 Vgl. hierzu *Jan Szezepanski: "Die biographische Methode"* in *René König* (Hrsg.), *op. cit.*, 3. Aufl., Band 4, S. 226-252; *Louis Gottschalk* u. Mitarb.: *The Use of Personal Documents in History, Anthropology and Sociology*, Social Science Research Council, New York 1945.
- 15 Dies ist ausführlicher beschrieben in *Erwin K. Scheuch: "Entwicklungsrichtungen bei der Analyse sozialwissenschaftlicher Daten"*, in *René König* (Hrsg.): *op. cit.*, 3. Aufl., S. 161 bis 238.
- 16 Die Bibliographie bei *H.J. Hummel: Probleme der Mehrebenenanalyse* Teubner, Stuttgart 1974 zeigt, daß selbst deutschsprachige Autoren zu diesen Themen durchweg in englischer Sprache veröffentlichten. Dies gilt beispielsweise für alle einschlägigen Schriften von

- E.K. Scheuch* zu diesem Thema. In Europa allgemein gab es für diese Themen in den 60er Jahren kein Publikum.
- 17 Die professionelle Öffentlichkeit für Methodenfragen der Sozialforschung war in den USA jedoch bis in die fünfziger Jahre hinein nicht so sehr die Soziologie i.e.S., sondern Anwendungsgebiete für Soziologie, insbesondere Sozialarbeit und Konsumforschung. Die ersten methodologischen Gesamtdarstellungen erschienen mit Unterstützung der Russel Sage Foundation, die besonders an einer Praxis-orientierten Soziologie und dem Beruf Sozialarbeit interessiert war. Vgl. hierzu *Pauline V. Young: Scientific Social Surveys and Research*, Prentice-Hall, New York 1939, einführende Kapitel. Für diese Orientierung der frühen Sozialforschung an der Praxis, ihre Verbindung zu Behörden und Reformbewegungen und das Bemühen um Standardisierung des Vorgehens siehe *Carl C. Taylor*: "The Social Survey, Its History and Methods", in: *University of Missouri Bulletin*, Nr. 28, 1919; und *Erle F. Young*: "The Social Base Map", in: *Journal of Applied Sociology*, Bd. 9 (1925), S. 202-206. Gerade die Verwendung von Personal, das weder Hochschul-lehrer war noch diese Laufbahn anstrebte und die Notwendigkeit, sich gegenüber zwar sach- aber nicht fachkundigen Instanzen zu erklären und zu legitimieren, beförderte die Ausbildung von Methodologie als erlernbare Technologie. Als Beleg für diesen Prozeß der Selbsterklärung mag dienen *Ernest W. Burgess*: "The Value of Sociological Community Studies for Work of Social Agencies", in: *Social Forces*, Bd. 8 (1930), S. 481-491.
 - 18 In den Vereinigten Staaten ist es für die führenden Zeitschriften üblich, die zur Publikation eingereichte Artikel nach Unkenntlichmachung Gutachtern zuzusenden, die teilweise wie die Gutachter der DFG von der Profession gewählt werden; dies sind die sogenannten "reviewed journals". Die durchschnittlich Rate der Zurückweisung liegt in den "klassischen" Naturwissenschaften bei 30% und weniger, bei der Soziologie um 78%, der politischen Wissenschaft sogar bei 84%; cf. *Lowell L. Hargens: Patterns of Scientific Research*, American Sociological Association, Washington 1975, S. 19, Tabelle 2-5. Inzwischen ist für die Soziologie die Rate der Zurückweisung wegen mangelnder Qualität für die wichtigsten Zeitschriften auf nahezu 90% gestiegen – ein Indiz für den Rückgang professioneller Qualifikation im Gefolge der Veränderung von Standards für Anstellung an amerikanischen Universitäten.
 - 19 Diese Schlußfolgerung wird bekräftigt durch eigene Lehrtätigkeit an einer Reihe amerikanischer Universitäten. Sie wird bestätigt durch die Erfahrungen mit dem Unterrichten amerikanischer Studenten in der Bundesrepublik und durch Beobachtungen während internationaler Ausbildungsseminare. In einer neuen Erhebung erklärten 48% der Soziologen, die "graduate" Studenten (in etw: Doktoranden) unterrichteten, sie seien mit dem Zustand der Ausbildung unzufrieden; nach *Lowell L. Hargens, op. cit.*, S. 71, Tabelle 4-1.
 - 20 Je höher die Qualifikation eines Wissenschaftlers, desto weniger ist er im Regelfalle an Forschung beteiligt, formuliert *Burkart Lutz* aufgrund seiner Erhebung über die Lage der Sozialforschung für die DGS; *Burkart Lutz*: "Zur Lage der soziologischen Forschung in der Bundesrepublik", zitiert nach dem Originalbericht, S. 44. Ähnlich im gleichen Manuskript auf S. 55, wo festgestellt wird, daß in der nicht-kommerziellen Forschung die Anfänger dominieren.
 - 21 *Eberhard Erbslöh* u. Mitarb.: *Studien zum Interview*, Anton Hain, Meisenheim 1973; *Hartmut Esser: Soziale Regelmäßigkeiten des Befragtenverhaltens*, Anton Hain, Meisenheim 1975.
 - 22 *Wilfried Becker: Beobachtungsverfahren in der demoskopischen Marktforschung*, Eugen Ulmer, Stuttgart 1973; *Jürgen Friedrichs und Hartmut Lüdtke: Teilnehmende Beobachtung*, Beltz, Weinheim, 2. Aufl. 1973; *Karl-Wilhelm Grümer: Beobachtung*, Teubner, Stuttgart 1974.
 - 23 *Dieter Holtmann*: "Multidimensionale Skalierung, Methode und ihre Anwendung in den Sozialwissenschaften", unveröffentlichte Dissertation Köln 1974; *Erwin K. Scheuch und Helmut Zehnpfenning*: "Skalierungsverfahren in der Sozialforschung", in *René König* (Hrsg.): *op. cit.* Band 3a, 1974. Eine sehr eigenständige Anwendung bringt *Wolfgang Manz: Das Stereotyp*, Anton Hain, Meisenheim 1969.
 - 24 *Wolfgang Sodeur: Empirische Verfahren zur Klassifikation*, Teubner, Stuttgart 1974; *Theodor Harder: Dynamische Modelle in der empirischen Sozialforschung*, Teubner, Stuttgart 1973; in Passagen auch *W. Bungard* und *H.E. Lück: Forschungsartefakte und nicht-reaktive Meßverfahren*, Teubner, Stuttgart 1974.

- 25 *H.J. Hummell: Probleme der Mehrebenenanalyse*, Teubner, Stuttgart 1973; *Steffen Harbordt: Computersimulation in den Sozialwissenschaften*, Rowohlt, Hamburg 1974, 2 Bände; *Klaus Allerbeck: Datenverarbeitung in der empirischen Sozialforschung*, Teubner, Stuttgart 1972; *Jürgen Kriz: Datenverarbeitung für Sozialwissenschaftler*, Rowohlt, Hamburg 1975.
- 26 Es handelt sich um die von *E.K. Scheuch* und *H. Sahner* herausgegebene Reihe "Studien-skripten zur Soziologie", Teubner-Verlag, Stuttgart; anders konzipiert aber mit gleicher Zielgruppe ist die Reihe "Sozialwissenschaft" innerhalb von "rororo studium", Rowohlt Hamburg, die von *F.X. Kaufmann* und *J. Matthes* herausgegeben wird. Ambitionierter ist die von *J. v. Koolwijk* und *M. Wieken-Mayser* herausgegebene Monographie-Reihe "Techniken der empirischen Sozialforschung", Oldenbourg, München.
- 27 *René König* (Hrsg.): *Handbuch der empirischen Sozialforschung*, Enke, Stuttgart, dessen 3. Auflage seit 1973 im Erscheinen begriffen ist. Die im engeren Sinne methodischen Aufsätze liegen bereits überarbeitet vor.
- 28 *Gardner, Lindzey* (Hrsg.): *Handbook of Social Psychology*, Addison Wesley, Cambridge (Mass), 2. Aufl. 1969.
- 29 Man vergleiche die heutige Situation mit der bis etwa Mitte der sechziger Jahre. Über einen Zeitraum von über 12 Jahren dominierten als Einführung die beiden Bände der "Praktischen Sozialforschung": *René König* mit *Dietrich Rüschemeyer* und *Erwin K. Scheuch* (Hrsg.): *Das Interview*, 1. Aufl. 1952 und *René König* mit *Peter R. Heintz* und *Erwin K. Scheuch* (Hrsg.): *Beobachtung und Experiment*, 1. Aufl. 1956, beide in Verlag für Politik und Wirtschaft, später Kiepenheuer & Witsch, Köln. Obgleich diese Bücher in mehreren Auflagen erschienen, waren diese Auflagen gering. Dann erschienen während der "Kulturrevolution", deren Themen ja auch eine prinzipielle Kritik der "positivistischen" Sozialforschung einschlossen, in geringem zeitlichen Abstand verschiedene methodologische Einführungen, die nacheinander rasch verschiedene Auflagen erlebten; die Auflagehöhe überstieg jeweils bei weitem die der beiden Bände von *René König*. Von diesen Einführungen seien erwähnt: *Renate Mayntz*, *Kurt Holm* und *Peter Hübner: Einführung in die Methoden der empirischen Soziologie*, Westdeutscher Verlag, 1. Aufl. 1969, Köln; *Peter Atteslander: Methoden der empirischen Sozialforschung*, Walter de Gruyter, Berlin, 1. Aufl. 1969 (bis heute 4 Auflagen!); *Jürgen Friedrichs: Methoden der empirischen Sozialforschung*, Rowohlt, Reinbeck bei Hamburg 1973. Das Buch von *Elisabeth Noelle-Neumann: Umfragen in der Massengesellschaft*, Rowohlt, Reinbeck bei Hamburg, zuerst 1963, behandelt zwar nur eine allerdings besonders wichtige Verfahrensweise, die Umfrageforschung, wurde aber dennoch später zu einem "Bestseller". Insofern die Auflagehöhe als Indiz gelten kann, war selbst während der "Kulturrevolution" das Interesse von Studenten der Sozialwissenschaften an der Technik der Sozialforschung als Grundlage der Profession groß.
- 30 Der "Positivismusstreit" beginnt erst jetzt seine Eigenschaft als deutscher Evergreen zu verlieren. Seinen wissenschaftlich relativ respektabelsten Niederschlag fand er in der zu Beginn der "Kulturrevolution" neu gegründeten Zeitschrift *Quality and Quantity*, 1. Jahrgang 1967, Società editrice il Mulino, Bologna. Diese Stilisierung in zwei Bekenntnis-Lager hatte weder mit der empirischen Forschung nennenswert etwas zu tun, noch mit den tatsächlichen methodischen Positionen der insgesamt als "Positivisten" apostrophierten Personen. Der Publikumserfolg dieser überflüssigen Diskussion ist nicht aus der Soziologie als einer sozialwissenschaftlichen Disziplin zu erklären, sondern aus Usancen des Feuilletons und den damit verknüpften literarischen Kontroversen. Werden unterschiedliche Positionen zu antagonistischen Schulen aufbereitet – auch wenn sie in Wirklichkeit unterschiedliche Zielsetzungen und damit wenig miteinander zu tun haben – so ist ein Leitmotiv bestimmt, durch das ein sich selbst befördernder Streit möglich wird. Erst durch die Aufbereitung eines Leitmotivs, für das der Charakter der Vorlagen allerdings wichtig ist, wird eine Einigung von Publikum und Autoren auf ein Thema möglich, wird die normale Aufspaltung in eine Vielzahl von Kleintemen aufgehoben – genau analog zur Unterhaltungsindustrie mit ihren Abfolgen von Moden. Hierfür eignen sich vor allem Vorlagen, die sich auf eine Formel bringen lassen, die sich auf weitgehend beliebig aber sehr viele Sachverhalten anwenden läßt, und die möglichst mit einer bekenntnisthaften Entscheidung verknüpft ist. Mit diesen Eigenschaften ist auch die zweit-erfolgreichste dieser aufbereiteten Kontroversen zu erklären, die Erörterung von "Systemtheorie" und der

- neueren Version "emanzipatorischer" Soziologie, *ulgo* der Streit zwischen *Habermas* und *Luhmann*. Für diese aufbereitete Kontroverse ist es unerheblich, zu welchem Grade *Luhmann* wirklich das repräsentiert, was "Systemtheorie", insbesondere außerhalb der Sozialwissenschaften, wirklich ist; wichtiger ist eine präzise Formulierung einfacher Gedanken – etwa "Reduktion von Komplexität" statt "Vereinfachung". Mit diesen Hinweisen soll die Behauptung verständlicher werden, daß der Positivismusstreit und die ideologischen Auseinandersetzungen über Sozialforschung einer Nachfrage bei einem allgemeinen Publikum entgegenkamen und nicht Ausdruck einer wissenschaftsinternen Entwicklung war, also auch nur begrenzt auf die Sozialforschung zurückwirken konnte.
- 31 Dies ergab eine Durchsicht der folgenden Zeitschriften: *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *Soziale Welt*, *Zeitschrift für Soziologie*, *Kyklos*, *Zeitschrift für die gesamte Staatswissenschaft*, *Hamburger Jahrbuch für Wirtschafts- und Gesellschaftspolitik*, *Schmoller's Jahrbuch*, *Verfassung und Verfassungswirklichkeit*, *Sozialwissenschaftliches Jahrbuch für Politik*. Am ehesten sind noch die "Kölner Zeitschrift" und die "Zeitschrift für Soziologie" hier relevante Periodika; wir entschieden uns schließlich, die "Kölner Zeitschrift" als breit eingeführte Quelle in die Übersicht einzubeziehen, die noch weniger allgemein gelesene "Zeitschrift für Soziologie" aber nicht mehr zu berücksichtigen. Die fehlende allgemeine Verbreitung war auch der Gesichtspunkt, "Quality and Quantity" nicht mehr auszuwerten, obgleich hier nicht selten Aufsätze erscheinen, die für die Methodologie der Sozialforschung interessant sind; in den Themen werden allerdings die Aufsätze in den allgemein von Sozialforschern gelesenen Zeitschriften weitestgehend wiederholt, sodaß sich der Eindruck ergibt, es handele sich um eine "Ausweich"-Quelle für die generell eingeführten Publikationen.
- 32 Die Auswertung wurde vorbereitet durch Dr. *Ekkehart Zimmermann*, jetzt Gesamthochschule Wuppertal und damals wissenschaftlicher Assistent am Institut für angewandte Sozialforschung in Köln. Ihm sei für diese anfängliche Mithilfe gedankt.
- 33 Die "Kölner Zeitschrift" ist zwar von den deutschsprachigen Periodika für Soziologie mit Abstand international am weitesten verbreitet, ist aber dennoch im wesentlichen nur eine bundesrepublikanische Veröffentlichung. Die anderen in die Untersuchung aufgenommen Periodika können jedoch als internationale Foren für das gelten, was auch im Deutschen "state of the art" genannt wird.
- 34 Diese Aussage über die Funktionsunterschiede zwischen Readern und Aufsätzen in Fachzeitschriften – kodifizierter und für Lehre geeigneter Teil der methodologischen Literatur versus fortgeschrittenster Teil methodologischer Diskussion – gilt heute nur noch eingeschränkter als vor etwa 10 Jahren. Die von *Blalock* herausgegebenen Bücher sind hierfür das wichtigste Beispiel; z.B. *Hubert M. Blalock* und *Ann B. Blalock* (Hrsg.): *Methodology in Social Research*, McGraw-Hill, New York 1968. Vor allem in Fremdsprachen sind Reader und Anthologien wichtigere Medien, gerade weil hier die Zeitschriften nicht die zentrale Bedeutung für die Soziologie haben, wie dies für die USA zutrifft. Wichtige Beispiele hierfür sind das wiederholt erwähnte Werk von *René König* (Hrsg.): "Handbuch der empirischen Sozialforschung", oder *Raymond Boudon* und *Paul Lazarsfeld* (Hrsg.): *L'analyse empirique de la causalité*, Mouton, Paris 1966 und *Francois Chazel*, *Raymond Boudon* und *Paul Lazarsfeld* (Hrsg.): *L'analyse des processus sociaux*, Mouton, Paris 1970.
- 35 Es ist in den Vereinigten Staaten üblich, daß immer wieder einzelne Artikel themensetzend wirken. In dieser Eigenheit, daß sich mit großer Schnelligkeit sehr viele Autoren auf ein Thema konzentrieren, kann ein Vorzug des dortigen Publikationswesens gesehen werden: Auf diese Weise werden Themen wirklich durchdiskutiert. Es ist nicht schwierig, unter rein intellektuellen Gesichtspunkten diese Art der Konzentration abzuwerten: Im Nachhinein erweisen sich viele Themen als überschätzt, Themen werden wie Modeartikel fallen gelassen, oder die Beschäftigung wird zu einer Spezialität ohne weitere Ausstrahlung. Beispiele sind aus den fünfziger Jahren die Art der Beschäftigung mit "Status-Inkonsistenz", oder das Problemverständnis beim Thema "eheliche Machtverhältnisse", oder das ziemlich ergebnislose Auslaufen der Beschäftigung mit der funktionalistischen Schichtungstheorie oder der Art, sozialen Status zu messen. Dennoch werden auf eben diesem Wege Themen zu Angelegenheiten der ganzen Profession, was für Aufsätze in deutschsprachigen Zeitschriften fast nie (bemerkenswerte Ausnahme: der Positivismusstreit zwischen *Habermas* und *Albert* in der *Kölner Zeitschrift*) der Fall ist. Wir erhalten unsere

Themen auf dem Umweg über die USA und damit auch als Konsequenz die "Modethemen".

- 36 In der amerikanischen Soziologie ist es zur Tradition geworden, bei den periodischen Bestandsaufnahmen die Zitierungshäufigkeit und Titelzählungen als empirische Grundlage für Trenddiagnosen zu nehmen. Bevorzugtes Publikationsorgan für diese Bestandsaufnahmen ist die seit 1965 erscheinende Zeitschrift *American Sociologist*, gelegentlich auch das Nachrichtenblatt *Footnotes* der American Sociological Association. Die Argumente, die hier gegen eine vorwiegend zählende Bestandsaufnahme vorgebracht wurden, verlieren bei Betrachtungen an Bedeutung, die sich über längere Zeiträume als diese hier erstrecken.
- 37 Robert McGinnis: "Methods of Research – The New Developments", in *Talcott Parsons* (Hrsg.): *American Sociology*, Basic Books, New York 1968, S. 109.
- 38 *ibid.*, S. 107.
- 39 So in *Erwin K. Scheuch*: "Methodische Probleme gesamtgesellschaftlicher Analyse" in *Theodor W. Adorno* (Hrsg.): *Spätkapitalismus oder Industriegesellschaft?* Enke, Stuttgart 1969, S. 153-182. Es muß relativierend zu den hier zitierten beiden Standpunkten, *McGinnis* und *Scheuch*, denen die Erwartung eines raschen Fortschritts in der Forschungstechnologie und/oder Methodologie gemein ist, die nahezu gleichzeitige Beurteilung der Lage durch einen der bedeutendsten Methodologen und Förderer der mathematischen Statistik gegenüber gestellt werden. "The recent investment in mathematical models has generated little or no new theory or knowledge, and the arguments for using mathematical models to generate theories appear to have little merit . . . Although the concern with mathematical models has persisted, a greater awareness of their limitations has developed . . . An interest in mathematical training for sociologists persists, but the vigor with which this interest was once pursued appears to have waned," cf. *Edgar F. Borgatta*: "The Current Status of Methodology in Sociology", in *Borgatta* (Hrsg.): *Sociological Methodology 1969*, Jossey-Bass, San Francisco 1969, S.X. und XI. *Borgatta* ist auch skeptisch in seinem Urteil über die Bedeutung der Entwicklung von Hochleistungscomputern: "Indeed, it is doubtful that any of the machine advances that have occurred since 1960 have been of material relevance for social scientists, outside data analysis." *ibid.*, S. XIII. Insgesamt sieht *Borgatta* keine Anzeichen für eine besonders stürmische Entwicklung in der Sozialforschung.
- 40 Dennoch gibt es inzwischen Beispiele mathematischer Analysen von Elementarphänomenen, wie in *James S. Coleman*: *The Adolescent Society*, Free Press, Glencoe (Ill) 1961; oder *Rolf Ziegler*: *Kommunikationsstruktur und Leistung sozialer Systeme*, Anton Hain, Meisenheim 1968. Allerdings ist erst in Einzelfällen absehbar, wie der Schritt von der Analyse von Elementarphänomenen zu komplexen Phänomenen gelingen soll – am ehesten noch bei *Edward O. Laumann*: *Bonds of Pluralism*, John Wiley, New York 1973. Siehe ferner zur Charakterisierung dieser Forschungsrichtung *Raymond Boudon*: *Mathematische Modelle und Methoden*, Ullstein, Frankfurt 1972; *Raymond Boudon*: *L'analyse mathématique des faits sociaux*, Plon, Paris 1967; *Rolf Ziegler*: "Anwendung mathematischer Verfahren zur Analyse des Statuszuweisungsprozesses", vervielfältigtes Manuskript, Institut für Soziologie, Kiel 1975.
- 41 Diese Mathematisierung der Forschung zur horizontalen Mobilität ist verknüpft insbesondere mit den Namen *Samuel Stouffer*, *George K. Zipf* und *James Beshers*. Zur Einführung siehe: *Samuel Stouffer*: *Social Research to Test Ideas*, Free Press, Glencoe (Ill), 1962, S. 68-112.
- 42 *Erwin K. Scheuch*: "Entwicklungsrichtungen bei der Analyse sozialwissenschaftlicher Daten", in *René König* (Hrsg.), *op. cit.*, 3. Auflage 1973, Band I, S. 655.
- 43 "Employing Nominal Variables, Induced Variables, and Block Variables in Path Analysis", *Sociological Methods and Research*, Band 1, 1972, S. 147 ff.
- 44 "Indirect Effects in Path Analysis", *Sociological Methods and Research*, Band 1, 1972, S. 175 ff.
- 45 *Charles E. Werts*, *Karl G. Jöreskog* und *Robert L. Linn*: "Identification and Estimation in Path Analysis with Unmeasured Variables", *American Journal of Sociology*, Band 78 1972/73, S. 1469 ff.
- 46 *Erich Weede*: "Zur Methodik der kausalen Abhängigkeitsanalyse (Pfadanalyse) in der nicht-experimentellen Forschung", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*,

- Band 22, 1970, S. 532 ff.: *Erich Weede*: "Zur Pfadanalyse", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 24, 1972, S. 101 ff.
- 47 "The Multivariate Analysis of Ordinal Measures", *American Journal of Sociology*, Band 76, 1970/1, S. 908 ff.
- 48 *Henri Theil*: "On the Estimation of Relationships Involving Qualitative Variables", *American Journal of Sociology*, Band 76, 1970/1, S. 103 ff.; *Jae-On Kim*: "Predictive Measures of Ordinal Association", *American Journal of Sociology*, Band 76, 1970/1, S. 891 ff.
- 49 "Linking Form of Hypothesis to Type of Statistic – An Application of Goodman's Z", *American Sociological Review*, Band 37, 1972, S. 357 ff.
- 50 *Robert P. Althaus* und *Michael Wigler*: "Standardization and Component Analysis", *Sociological Methods and Research*, Band 1, 1972, S. 1971 ff.
- 51 *David D. McFarland*: "Intragenerational Social Mobility as a Markov Process – Including a Time-Stationary Markovian Model that Explains Observed Declines in Mobility Rates Over Time", *American Sociological Review*, Band 35, 1970, S. 463 ff.; *Seymour Spilerman*: "The Analysis of Mobility Processes by the Introduction of Independent Variables into a Markov Chain", *American Sociological Review*, Band 37, 1972, S. 277 ff.; *Seymour Spilerman*: "Extensions of the Mover-Stayer Model", *American Journal of Sociology*, Band 78, 1972/73, S. 599 ff.; *Gudmund Hernes*: "A Markovian Approach to Measures of Association", *American Journal of Sociology*, Band 75, 1969/70, S. 992 f.
- 52 *David R. Heise*: "Separating Reliability and Stability in Test-Retest Correlation", *American Sociological Review*, Band 34, 1969, S. 93 ff.
- 53 *J. David Martin* und *Louis N. Gray*: "Measurement of Relative Variation", *American Sociological Review*, Band 36, 1971, S. 496 ff.
- 54 "Duncan Requests Reconsideration of Award", *ASA Footnotes*, Dezember 1974, S. 2. Es handelt sich um den Abdruck eines Leserbriefes, in dem *Duncan* die Gewährung des *Samuel Stouffer* Preises an ihn und *Leo Goodman* beanstandet und für *Goodman* als alleinigen Preisträger plädiert.
- 55 eodem loco.
- 56 *Leo A. Goodman*: "A General Model for the Analysis of Surveys", *American Journal of Sociology*, Band 77, 1971/72, S. 1035 ff.
- 57 *Leo A. Goodman*: "How to Ransack Social Mobility Tables and Other Kind of Cross-Classification Tables", *American Journal of Sociology*, Band 75, 1969, S. 1 ff.
- 58 *Leo A. Goodman*: "Causal Analysis of Data from Panel Studies and Other Kinds of Surveys", *American Journal of Sociology*, Band 78, 1972/73, S. 1135.
- 59 *Leo A. Goodman*: "The Analysis of Systems of Qualitative Variables When Some of the Variables Are Unobservable", *American Journal of Sociology*, Band 79, 1973/74, S. 1179 ff.
- 60 *Leo A. Goodman*: "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables", *American Sociological Review*, Band 37, 1972, S. 28 ff.
- 61 *Jonathan Cloud* und *Graham M. Vaughan*: "Using Balanced Scales to Control Acquiescence", *Sociometry*, Band 33, 1970, S. 193 ff.
- 62 "The Srole Items and Acquiescence", *American Sociological Review*, Band 35, 1970, S. 287 ff.
- 63 *Derek L. Phillips* und *Kevin J. Clancy*: "Some Effects of 'Social Desirability' in Survey Studies", *American Journal of Sociology*, Band 77, 1971/72, S. 921 ff.
- 64 *George W. Bohrnstedt*: "A Quick Method for Determining the Reliability and Validity of Multiple-Item Scales", *American Sociological Review*, Band 34, 1969, S. 542.
- 65 "A Review and Comparison of Simple Statistical Tests for Scalogram Analysis", *American Sociological Review*, Band 38, 1973, S. 238 ff.
- 66 "A Note of Extreme Caution on the Use of Guttman Scales", *American Journal of Sociology*, Band 74, 1968/69, S. 296 ff.
- 67 *Kurt Holm*: "Zuverlässigkeit von Skalen und Indizes", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 22, 1970, S. 356 ff.; *Kurt Holm*: "Gültigkeit von Skalen und Indizes", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 22, 1970, S. 693 ff.
- 68 Ich habe nur einen wesentlichen Einwand gegen einen Aspekt des ersten der beiden Aufsätze: Die Unterstellung von *Holm*, daß er keinen nennenswerten Unterschied zwischen Skalen und Indizes sehe, *ibid.*, S. 356 und 385. Ich selbst sehe als entscheidendes Kriterium für Skalen ein restriktives Modell, das an den Daten scheitern kann; Indizes "stim-

- men" zunächst immer. Allerdings ist *Holm* zuzustimmen, daß für die Arbeitsgänge der Prüfung von Zuverlässigkeit und Gültigkeit ein solcher Unterschied keine Rolle spielt.
- 69 *Philip M. Lankford*: "Comparative Analysis of Clique Identification Methods" *Sociometry*, Band 37, 1974, S. 287-305.
- 70 *Edmund R. Peay*: "Hierarchical Clique Structures", *Sociometry*, Band 37, 1974, S. 54 ff.
- 71 *Paul W. Holland* und *Samuel Leinhardt*: "A Method for Detecting Structure in Sociometric Data", *American Journal of Sociology*, Band 76, 1970/71, S. 492 ff.
- 72 *John T. Toby*: "A Measure of Sociometric Reciprocity", *Sociological Methods and Research*, Band 1, 1973, S. 497 ff.
- 73 Der damalige Stand der Diskussion wird ersichtlich aus *J.L. Moreno* (Hrsg.): *The Sociometry Reader*, Free Press, Glencoe (Ill), 1960. Wahrscheinlich war es für die Entwicklung der Soziometrie Ende der sechziger Jahre auch wichtig, daß die von *Moreno* entfachte unnötige Kontroverse über "heiße" (anwendungsbezogene im Sinne von Therapie) und "kalte" (d.h. Nutzung der Werte nur für Zwecke der Forschung) Soziometrie schlicht vergessen wurde.
- 74 Zur Einführung in die Problematik nach Überwindung der ersten Kontroversen in der Form eines Schulenstreits zwischen Aggregatdaten-Analitikern und ihren Kritikern siehe *Erwin K. Scheuch*: "Cross-National Comparisons Using Aggregate Data; Some Substantive and Methodological Problems", in *Richard L. Merritt* und *Stein Rokkan* (Hrsg.): *Comparing Nations – The Use of Quantitative Data in Cross-National Research*, Yale University Press, New Haven 1966; *Erwin K. Scheuch*: "Ökologischer Fehlschluß", in *Wilhelm Bernsdorf* (Hrsg.): *Wörterbuch der Soziologie*, Enke, 2. Auflage, Stuttgart 1969.
- 75 *John L. Hammond*: Two Sources of Error in Ecological Correlations", *American Sociological Review*, Band 38, 1973, S. 764 ff.
- 76 *Gudmund R. Iversen*: "Recovering Individual Data in the Presence of Group and Individual Effects", *American Journal of Sociology*, Band 79, 1973/74, S. 420 ff.
- 77 *Michael T. Hannan* und *Leigh Burstein*: "Estimation From Grouped Observations", *American Sociological Review*, Band 39, 1974, S. 374 ff.
- 78 *Robert M. Hauser*: Context and Conseq – A Cautionary Tale", *American Journal of Sociology*, Band 75, 1969/70, S. 645 ff.
- 79 Eine differenzierte Betrachtung der Messniveaus findet sich in *Johann Pfanzagl*: *Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens*, Physika, Würzburg 1959 und *Johann Pfanzagl*, *V. Baumann* und *H. Hubert*: *Theory of Measurement*, Würzburg 1968.
- 80 Eine Übersicht über die Entwicklung dieses Schulenstreits und seines heutigen Standes findet sich in *Alan C. Acock* und *J. David Martin*: "The Undermeasurement Controversy – Should Ordinal Data be Treated as Interval?", *Sociology and Social Research*, Band 58, 1974, S. 427 ff.
- 81 Als Beispiel *ibid.*, S. 430/1.
- 82 *Sanford Labovitz*: "Statistical Usage in Sociology – Sacred Cows and Ritual", *Sociological Methods and Research*, Band 1, 1972, S. 13 ff.; *Richard P. Boyle*: "Path Analysis and Ordinal Data", *American Sociological Review*, Band 38, 1973, S. 461 ff.; *Sanford Labovitz*: "The Assignment of Numbers to Rank Order Categories", *American Sociological Review*, Band 35, 1970, S. 515 ff.
- 83 *A.A. Hunter*: "On the Validity of Measures of Association – The Nominal-Nominal Two-By-Two Case", *American Journal of Sociology*, Band 79, 1973,
- 84 *Herbert L. Costner*: "Theory, Deduction, and Rules of Correspondence", *American Journal of Sociology*, Band 75, 1969, S. 245 ff.; *Hubert M. Blalock*: "Multiple Indicators and the Causal Approach to Measurement Error", *American Journal of Sociology*, Band 75, 1969, S. 264 ff.
- 85 *Friedrich Lösel* und *Werner Wüstendörfer*: "Zum Problem unvollständiger Datenmatrizen in der empirischen Sozialforschung", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 26, 1974, S. 342 ff.
- 86 *Richard A. Zeller* und *Zachary H. Levine*: "The Effects of Violating the Normality Assumption Underlying r", *Sociological Methods and Research*, Band 2, 1974, S. 511 ff.; *Marilyn B. Brewer*, *Donald T. Campbell* und *William D. Crano*: "Testing a Single-Factor Model as an Alternative to the Misuse of Partial Correlations in Hypothesis-testing in Research", *Sociometry*, Band 33, 1970, S.1 ff.

- 87 David L. Klemmack, Thomas A. Leggette und Lawrence S. Mayer: "Non-Random Exogenous Variables in Path Analysis", *American Sociological Review* Band 38, 1973, S. 778 ff.
- 88 Kathleen S. Crittenden und Richard J. Hill: "Coding Reliability and Validity of Interview Data", *American Sociological Review*, Band 36, 1970, S. 1073 ff.
- 89 Für die empirische Untersuchung von Jürgen van Koolwijk: "Unangenehme Fragen – Paradigma für die Reaktionen des Befragten im Interview", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 21, 1969, S. 864 ff. gibt es kein zureichendes Äquivalent in englischsprachigen Veröffentlichungen; dies ist ein wichtiger Beitrag zum Wissen über das Interview. Wichtige Beiträge sind auch Kurt Holm "Theorie der Frage" und "Theorie der Fragebatterie", beides *Kölner Zeitschrift für Soziologie*, Band 26, 1974, S. 91 ff. und S. 317 ff.
- 90 Ansgar Weyman: "Bedeutungsfeldanalyse – Versuch eines neuen Verfahrens der Inhaltsanalyse am Beispiel Didaktik", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 25, 1973, S. 761 ff.
- 91 Siehe als Beleg Seymour Sudman und Norman M. Bradburn: *Response Effects in Surveys*, Aldine, Chicago 1974.
- 92 Sam D. Sieber: "The Integration of Fieldwork and Survey Methods", *American Journal of Sociology*, Band 78, 1972, S. 1335 ff.; der Titel des Aufsatzes ist übrigens ziemlich irreführend und der Aufsatz selbst nicht sehr gewichtig.
- 93 Daniel S. Claster und Howard Schwartz: "Strategies of Participation in Participant Observation", *Sociological Methods and Research*, Band 1, 1972, S. 65 ff.
- 94 Richard A. Zeller und Richard B. Warnecke: "The Utility of Intervening Constructs in Experiments", *Sociological Methods and Research*, Band 2, 1973, S. 85 ff.; Bruce C. Straits, Paul L. Wuebben und Theophile J. Majka: "Influences on Subjects' Perceptions of Experimental Research Situations", *Sociometry*, Band 35, 1972, S. 499; J. David Martin: "Suspicion and the Experimental Confederate – A Study of Role and Credibility", *Sociometry*, Band 33, 1970, S. 178 ff.
- 95 Als Beispiele siehe Roland K. Hawkes: "Some Methodological Problems in Explaining Social Mobility", *American Sociological Review*, Band 37, 1972, S. 294 ff. und Karl Ulrich Mayer und Walter Müller: "Trendanalyse in der Mobilitätsforschung", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Band 23, 1971, S. 761 ff.
- 96 Alan C. Acock und Melvin L. Defleur: "A Configurational Approach to Contingent Consistency in the Attitude-Behavior Relationship", *American Sociological Review*, Band 37, 1972, S. 714-726.
- 97 Murray Webster: "Psychological Reductionism, Methodological Individualism, and Large-Scale Problems", *American Sociological Review*, Band 38, 1973, S. 258-273.
- 98 Kenneth D. Bailey: "Monothetic and Polythetic Typologies and Their Relation to Conceptualization, Measurement and Scaling", *American Sociological Review*, Band 38, 1973, S. 18-33.
- 99 Wenngleich für das Public Opinion Quarterly die Artikel als Zählheiten einander ähnlicher sind als dies für die anderen Zeitschriften galt, so verbleibt doch eine zweite, entscheidende Problematik, welche die Aussagekraft solcher Literaturzählungen einschränkt: Die Artikel behandeln nicht selten mehr als ein Thema. Beispiel: Soll ein Beitrag wie der von Arnold S. Linsky: "Stimulating Responses to Mailed Questionnaires", *Public Opinion Quarterly*, Band 39, 1975, S. 82 ff. eher der Kategorie "Ausschöpfung" oder "Schriftliche Befragungen" zugeordnet werden? In diesem Falle wurde die Kategorie "Ausschöpfung" bevorzugt, weil dies in der Kombination der beiden Kategorien der eingrenzendere Begriff ist. Offensichtlicher Nachteil: für die Besetzung der Kategorie "schriftliche Befragungen" wird damit ein Fall zu wenig ausgewiesen; Doppelzählungen wären aber noch problematischer gewesen. Insgesamt sind jedoch in der Übersicht die Schwerpunkte annähernd zutreffend wiedergegeben.
- 100 Peter H. Rossi und Robert Crain: "The NORC Permanent Community Sample", *Public Opinion Quarterly*, Band 32, S. 261 ff. Für die folgenden Angaben von Artikeln im Public Opinion Quarterly wird die Abkürzung POQ verwendet werden.
- 101 Mabel S. Sanderson: "The Usefulness of Data Provided by the Decennial Census", *POQ*, Band 33, 1969, S. 318 ff.

- 102 *Hans Zeisel*: "Opinion Polling in the Renaissance", *POQ*, Band 37, 1973, S. 282 ff.
- 103 *Jacob Jacob*: "The Construct Validity of Opinion Leadership", *POQ*, Band 1974, S. 81 ff.
- 104 Vergleiche hierzu die von *Leslie Kish*: *Survey Sampling*, John Wiley, New York 1965 als offen gekennzeichneten Probleme; cf. insbesondere S. 574-600. *Leslie Kish* ist als Leiter der Stichprobenabteilung des Survey Research Center der Universität Michigan einer der erfahrensten Praktiker und zugleich ein bedeutender Theoretiker der Anwendung von Stichprobenverfahren auf menschliche Bevölkerungen. Dieses Buch ist immer noch die wichtigste praxisnahe Darstellung für Stichprobenprobleme in der Umfrageforschung.
- 105 *Joseph B. Perry*: "A Note on the Use of Telephone Directories as a Sample Source," *POQ*, Band 32, 1968/69, S. 691 ff.; *Mathew Hauck* und *Michael Cox*: "Locating a Sample by Random Digital Dialing", *POQ*, Band 38, S. 253 ff.
- 106 *Joseph R. Hochstim* und *Demetrios A. Athanasopoulos*: "Personal Follow-Up in a Mail Survey – Its Contribution and Its Cost," *POQ*, Band 34, 1970, S. 69 ff.; *James C. Hacker* und *Patricia Bourgette*: "Dollars, Dissonance, and Survey Returns", *POQ*, Band 37, 1973, S. 276 ff.; *Wayne E. Hensley*: "Increasing Response Rate by Choice of Postage Stamps", *POQ*, Band 38, 1974, S. 280 ff.
- 107 Vgl. *Arnold S. Linsky*: "Stimulating Responses to Mailed Questionnaires", *POQ*, Band 39, 1975, S. 82 ff.; *George C. Mayers*: "The Elusive Male – Some Methodological Notes on Survey Research Design", *POQ*, Band 33, 1969, S. 255
- 108 *Barbara S.* und *Bruce P. Dohrenwend*: Sources of Refusals in Surveys", *POQ*, Band 32, 1968, S. 74 ff.; *Carol H. Fuller*: "Weighting to Adjust for Survey Nonresponse," *POQ*, Band 38, 1974, S. 247 ff.
- 109 Die repräsentativste Zusammenfassung von Originalforschungen ist immer noch *Herbert H. Hyman* u. Mitarb.: *Interviewing in Social Research*, University of Chicago Press, Chicago 1954, Siehe ferner *Seymour Sudman* und *Norman M. Bradburn*: *Response Effects in Surveys*, Aldine, Chicago 1974; *Mathew Hauck* und *Stanley Steinkamp*: *Survey Reliability and Interviewer Competence*, University of Illinois Press, Urbana (Ill) 1964; *Erwin K. Scheuch*: "Das Interview in der Sozialforschung", in *Renè König, op. cit.*, Band 2, 3. Aufl. 1973, S. 66-190.
- 110 *Elisabeth Noelle-Neumann*: "Wanted – Rules for Wording Structured Questionnaires", *POQ*, Band 34, 1970, S. 191; *Barbara S. Dohrenwend*, *John Colombotos* und *Bruce P. Dohrenwend*: "Social Distance and Interviewer Effects, *POQ*, Band 32, 1968, S. 410 ff.; *J. Allen Williams*: "Interviewer Role Performance – a Further Note on Bias in the Information Interview", *POQ*, Band 32, 1968, S. 287 ff.
- 111 *G. Allen Brunner* und *Stephen J. Carroll*: Weekday Evening Interviews of Employed Persons Are Better, *POQ*, Band 33, 1969, S. 265 ff.; *Joseph Zelan*: "Interviewing the Aged," *POQ*, Band 33, 1969, S. 420 ff.; *Edith Fein*: "Inner-City Interviewing", *POQ*, Band 34, 1970, S. 625 ff.
- 112 Vergleiche die in Anmerkung 21 angeführten Monographien.
- 113 *Bruce Dunning* und *Don Cahalan*: "By-mail vs. Field Self-Administered Questionnaires – An Armed Forces Survey," *POQ*, Band 37, 1973/74, S. 618 ff.
- 114 *Alan R. Andreasen*: "Personalizing Mail Questionnaire Correspondence", *POQ*, Band 34, 1970, S. 273 ff.; und *Edwin H. Carpenter*: "Personalizing Mail Surveys – a Replication and Reassessment", *POQ*, Band 38, 1974/75, S. 614 ff.; *Robert J. Parsons* und *Thomas Medford*: "The Effect of Advance Notice in Mail Surveys of Homogeneous Groups," *POQ*, Band 36, 1972, S. 258 ff.; *Robert E. Stevens*: "Does Precoding Questionnaires Affect Response Rates? ", *POQ*, Band 38, 1974/75, S. 621 ff.; *Richard A. Eisinger* u. Mitarb.: "Increasing Returns in International Surveys," *POQ*, Band 38, 1974/75, S. 124.; *Don A. Dillman*: "Increasing Mail Questionnaire Response in Large Samples of the General Public", *POQ*, Band 36, 1972, S. 254 ff.
- 115 *J. Scott Armstrong*: "Monetary Incentives in Mail Surveys", *POQ*, Band 39, 1975, S. 111 ff.
- 116 *Herbert H. Blumberg*, *Carolyn Fuller*, *A. Paul Hare*: "Response Rates in Postal Surveys", *POQ*, Band 38, 1974, S. 113 ff.
- 117 *Robert V. Stover* und *Walter J. Stone*: "Hand Delivery of Self-Administered Questionnaires", *POQ*, Band 38, 1974, S. 284 ff.; *Howard Schuman* und *Jean M. Converse*: "The Effects of Black and White Interviewers on Black Responses", *POQ*, Band 35, 1971, S. 44. Im letzteren Beitrag wird berichtet, daß die größten Abweichungen, die bei einer

- Untersuchung der Rassenunterschiede zwischen Interviewer und Befragten beobachtet wurden, die Bevorzugung von Unterhaltungsstars betrafen.
- 118 *Thomas Sheets*, u. Mitarb.: "Deceived Respondents – Once Bitten, Twice Shy", *POQ*, Band 38, 1974, S. 261. Die Beobachtung, daß Befragte, die sich getäuscht fühlten, in einem weiteren Interviewversuch negativ reagieren, ist in den USA heute allerdings schon von quantitativer Bedeutung. Zunehmend wirkt es sich als Störfaktor aus, daß alle Arten von Verkäufer und Sammlern die Sperre einer Wohnungstür durch ein fälschliches Auftreten als Interviewer zu überwinden versuchen.
- 119 *Robert G. Lehnen*: "Assessing Reliability in Sample Surveys", *POQ*, Band 35, 1971-72, S. 578 ff.; *Laurie A. Broedling*: "On More Reliability Employing the Concept of Reliability", *POQ*, Band 38, 1974, S. 372 ff.
- 120 *Joseph R. Hochstim* und *Karen S. Renne*: "Reliability of Response in a Sociomedical Population Study", *POQ*, Band 35, 1971, S. 69 ff. Wie häufig, so dürfte auch hier die Verlässlichkeit persönlicher Befragungen zu hoch ausgewiesen sein, weil die Abweichungen nicht bezogen auf die Möglichkeit zu Abweichungen ausgewiesen werden. Ansonsten wird wieder einmal belegt, daß "weiche" Daten weniger verlässlich als "harte" sind.
- 121 *Aage R. Clausen*: "Response Validity – Vote Report", *POQ*, Band 32, 1968/69, S. 588 ff.; *Charles N. Weaver* und *Carol L. Swanson*: "Validity of Reported Date of Birth, Salary, and Seniority", *POQ*, Band 38, 1974, S. 69 ff. In dem Aufsatz von *Clausen* dürfte allerdings die Gültigkeit der Angaben über Wahlverhalten übertrieben gut dargestellt sein. Ihm geht es übrigens nur um eine Prüfung, wie eine bessere Aggregatvorausage der Höhe der Wahlbeteiligung auf der Grundlage von Umfragen zu erreichen ist.
- 122 *Hugh J. Parry*, *Mitchell B. Balter* und *Ira H. Cisin*: "Primary Levels of Underreporting Psychotropic Drug Use", *POQ*, Band 34, 1970/71, S. 582 ff.
- 123 *Aage R. Clausen*: "Response Validity – Vote Report", *POQ*, Band 32, 1968, S. 588-606, u.a.S. 603 f.
- 124 *Eugene C. Hagburg*: "Validity of Questionnaire Data – Reported and Observed Attendance in an Adult Education Program", *POQ*, Band 32, 1968, S. 453.
- 125 *John A. Ballweg*: "Husband-Wife Response Similarities on Evaluative and Non-Evaluative Survey Questions", *POQ*, Band 33, 1969, S. 249-254.
- 126 *Carol H. Weiss*: "Validity of Welfare Mothers' Interview Responses", *POQ*, Band 32, 1968, S. 622-633.
- 127 Siehe hierzu auch *Erwin K. Scheuch*: "Das Interview in der Sozialforschung", in *René König* (Hrsg.): *op. cit.*, Band 2, S. 111-121. Dort werden auch weitere Ergebnisse von Originalerhebungen in der Bundesrepublik mitgeteilt.
- 128 Vergleiche *Herbert Hyman* u. Mitarb.: *Interviewing in Social Research*, University of Chicago Press, Chicago 1954, insbes. S.154 ff.
- 129 *Carol H. Weiss*, *op. cit.*, S. 629 f.
- 130 Diese Untersuchung wurde bisher nur zum Teil ausgewertet und bruchstückhaft veröffentlicht. Die wichtigsten Zeitschriftenaufsätze der ersten Runde der Analyse waren: *Hugh J. Parry* und *Helen M. Crossley*: "Validity of Responses to Survey Questions", *POQ*, Band 14, 1950, S. 61-80; *Helen M. Crossley* und *Raymond Fink*: "Response and Non-Response in a Probability Sample", *International Journal of Opinion and Attitude Research*, Band 5, 1951, S. 1-19; *J.J. Feldman*, *Herbert Hyman* und *Clyde W. Hart*: "A Field Study of Interviewer Effects on the Quality of Survey Data", *POQ*, 1951, S. 734-761.
- 131 *Don Cahalan*: "Correlates of Respondent Accuracy in the Denver Validity Survey", *POQ*, Band 32, 1968, S. 607-621.
- 132 *ibid.*, S. 610.
- 133 Es handelt sich um die "Community Chest", einen Zusammenschluß der privaten Wohltätigkeitsvereine auf lokaler Basis zum Zwecke der Spendensammlung. Es wird gewöhnlich starker Druck ausgeübt, entsprechend dem jeweiligen Status zu spenden. In diesem Widerspruch zwischen Freiwilligkeit im rechtlichen Sinne und standardisierten Erwartungen versuchen nicht wenige Amerikaner, heimlich den Erwartungen auszuweichen. Damit ist diese Frage ein gutes Kriterium für eine hohe Bereitschaft zu gültigen Antworten.
- 134 *Cahalan*, *op. cit.*, S. 621.
- 135 Siehe hierzu *W.H. Williams*: "The Systematic Bias Effects of Incomplete Responses in Rotation Samples," *POQ*, Band 33, 1969, S. 593-602.

- 136 Bruce K. Eckland: "Retrieving Mobile Cases in Longitudinal Surveys", *POQ*, Band 32, 1968, S. 51-64.
- 137 Norval D. Glenn: "Problems of Comparability in Trend Studies with Opinion Poll Data", *POQ*, Band 34, 1970, S. 81-91.
- 138 Shanto Iyengar: "Magnifying Relationships between Attitudinal Variables Using Panel Analysis", *POQ*, Band 38, 1974, S. 90-97; Robert G. Lehnen und Gary G. Koch: "Analyzing Panel Data with Uncontrolled Attrition", *POQ*, Band 38, 1974, S. 40-56.
- 139 Stanley Milgram, Leon Mann und Susan Harter: "The Lost Letter Technique – A Tool of Social Research", *POQ*, Band 29, 1965, S. 437 ff.
- 140 Allan B. Wicker: "A Failure to Validate the Lost-Letter Technique", *POQ*, Band 33, 1969, S. 260 ff.
- 141 Stanley Milgram: "Comment on 'A Failure to Validate the Lost-Letter Technique'", *POQ*, Band 33, 1969, S. 263 f.
- 142 D.M. Georgoff, B.J. Hersker und R.G. Murdick: "The Lost-Letter Technique – a Scaling Experiment", *POQ*, Band 36, 1972, S. 114-119.
- 143 Everett F. Cataldo u. Mitarb.: "Card Sorting as a Technique for Survey Interviewing", *POQ*, Band 34, 1970, S. 202 ff.
- 144 S. Stephen Kegeles, Sinton F. Fink und John P. Kirscht: "Interviewing a National Sample by Long-Distance Telephone," *POQ*, Band 33, 1969, S. 412-419
- 145 *ibid.*, S. 419.
- 146 Barbara Snell Dohrenwend: "An Experimental Study of Payments to Respondents," *POQ*, Band 34, 1970/71, S. 621-624.
- 147 Arnold S. Linsky: "Stimulating Responses to Mailed Questionnaires – A Review", *POQ*, Band 39, 1975, S. 82-101, insbesondere S. 99-100.
- 148 Stanford H. Odesky: "Localizing Early Vote Prediction Techniques", *POQ*, Band 33, 1969, S. 431 f.
- 149 Neal E. Cutler: "Generation, Maturation, and Party Affiliation – A Cohort Analysis", *POQ*, Band 33, 1969, S. 583-599.
- 150 John A. Sonquist: "Finding Variables that Work", *POQ*, Band 33, 1969, S. 83-95.
- 151 Edward R. Tufte: "A Note of Caution in Using Variables that have Common Elements", *POQ*, Band 33, 1969/70, S. 622 ff. Vergleiche auch Morris Rosenberg: "The Logical Status of Suppressor Variables", *POQ*, Band 37, 1973, S. 359-372.
- 152 Herbert H. Hyman: *Secondary Analysis of Sample Surveys – Principles, Procedures, and Potentialities*, John Wiley, New York 1972.
- 153 Siehe als Beispiele Herbert H. Hyman: "Black Matriarchy Reconsidered – Evidence from Secondary Analysis of Sample Surveys", *POQ*, Band 33, 1969, S. 346 ff.
- 154 David Nasatir: "The International Data Library and Reference Service", *POQ*, Band 32, 1968/69, S. 688 ff. Nasatir war damals Direktor der Datenbank in Berkeley, und der Aufsatz ist eine Werbeschrift für dieses Umfragearchiv – also auch dies keine zureichende Unterrichtung über Ressourcen für Sekundäranalysen.
- 155 Art Shulman: "A Comparison of Two Scales on Extremity Response Bias", *POQ*, Band 37, 1973, S. 407-412.
- 156 Alvin J. Silk: "Response Set and the Measurement of Self-Designated Opinion Leadership", *POQ*, Band 35, 1971/72, S. 383-397.
- 157 Thomas M. Ostrom: "Item Construction in Attitude Measurement", *POQ*, Band 35, 1971/72, S. 593-605.
- 158 Albert Chevam: "Minimum-Error Scalogram Analysis", *POQ*, Band 36, 1972/73, S. 379-387.
- 159 John P. Robinson: "Towards a More Appropriate Use of Guttman Scaling", *POQ*, Band 37, 1973/74, S. 260-267.
- 160 *Ibid.*, S. 266.
- 161 Karl E. Bauman und Charles L. Chase: "Interviewers as Coders of Occupation", *POQ*, Band 38, 1974, S. 107 ff.
- 162 G. Ray Funkhouser und Edwin B. Parker: "Analyzing Coding Reliability – The Random-Systematic-Coefficient", *POQ*, Band 32, 1968, S. 122-128.
- 163 Bruce Frisbie und Seymour Sudman: "The Use of Computers in Coding Free Responses", *POQ*, Band 32, 1968, S. 216-232. Das für diese Arbeiten zu Grunde gelegte Verfahren ist der "General Inquirer", dessen Prinzipien und Anwendungsformen dargestellt sind in

- Philip J. Stone u. Mitarb. (Hrsg.): *The General Inquirer – A Computer Approach to Content Analysis*, The MIT Press, Cambridge (Mass), 1966. Für eine erste, schnelle Information in deutscher Sprache siehe Erwin K. Scheuch und Ekkehard Mochmann: "Sozialforschung" in Peter Mertens: *Angewandte Informatik*, Walter de Gruyter, Berlin 1972, insbesondere S. 153 f.
- 164 "Aggregatdaten" meinte als Bezeichnung zunächst insbesondere die in amtlichen Zählungen, speziell Volkszählungen, gesammelten Daten, da diese Individualzählungen den Forschern nur in aggregierter Form zugänglich wurden. Die Bezeichnung "ökologische" Daten stellt darauf ab, daß die kleinste Aggregationsebene, für die Aggregatdaten den Forschern zugänglich werden, eine Gebietseinheit (Kreis, Stadt, Canton, etc) zu sein pflegt. Prozeßdaten ist als Oberbegriff gemeint, der Aggregatdaten und ökologische Daten, die ja als Folge regelmäßiger staatlicher Veranstaltungen entstehen, mit einschließt. Bisher ist übrigens die terminologische Verwirrung noch nicht wichtig geworden.
- 165 Als Einstieg in dieses Thema seien empfohlen: Charles L. Taylor (Hrsg.): *Aggregate Data Analysis*, Mouton, Paris 1968 sowie Mattei Dogan und Stein Rokkan (Hrsg.): *Quantitative Ecological Analysis in the Social Sciences*, MIT Press, Cambridge (Mass), 1969.
- 166 Vgl. Leslie D. Wilcox, Ralph M. Brooks, George M. Beal und Gerald Klomglang (Hrsg.): *Social Indicators and Societal Monitoring*, Elsevier, Amsterdam, 1972; dies ist die bei weitem beste Literaturübersicht. Siehe auch Eleanor Bernert Sheldon und Wilbert E. Moore (Hrsg.): *Indicators of Social Change*, Russell Sage, New York 1968 als u.E. bisher bester Reader. An deutschsprachigen Schriften seien erwähnt: *Arbeitskonferenz soziale Indikatoren*, I & II, Frankfurt 1972 und Mannheim 1973; Jörg Steinhausen: *Soziale Indikatoren*, Anton Hain, Meisenheim 1975. Als Beispiel für die Berichte, die aus dieser Entwicklung folgen, siehe Bundesministerium für innerdeutsche Beziehungen: *Materialien zum Bericht zur Lage der Nation, 1974*, Bonn 1974. Die erste bedeutende für Sozialwissenschaftler erstellte Sammlung war Bruce M. Russett u. Mitarb. (Hrsg.): *World Handbook of Political and Social Indicators*, Yale University Press, New Haven 1964; an diesem Band werden allerdings auch die methodischen Probleme einer Nutzung solcher Datensammlungen für Forschungszwecke sehr deutlich.
- 167 Aus einer sehr umfangreichen Literatur können hier Titel nur ausgewählt werden. Als Literaturübersichten sind besonders empfehlenswert: Stein Rokkan u. Mitarb. (Hrsg.): *Comparative Survey Analysis*, Mouton, Paris 1969, sowie Frederick W. Frey u. Mitarb. (Hrsg.): *Survey Research on Comparative Social Change*, MIT Press, Cambridge (Mass), 1969. Immer noch nützlich als Reader ist Stein Rokkan (Hrsg.): *Comparative Research Across Cultures and Nations*, Mouton, Paris 1969. Die beste systematische Darstellung dürfte immer noch sein Adam Przeworski und Henry Teune: *Theologic of Comparative Social Inquiry*, die ihr Thema am Beispiel eines international vergleichenden Projekts, der "Value Study" entwickeln. Gute Einsichten als Fallstudie der Probleme international vergleichender Untersuchungen finden sich in dem Bericht über die bisher umfangreichste international vergleichende Forschung, die 12-Länder-Studie über Zeitbudgets; Alexander Szalai u. Mitarb. (Hrsg.): *The Use of Time*, Mouton, den Haag, 1972.
- 168 Dies zeigt an den Daten der international vergleichende Erhebung "Die Welt im Jahr 2000" Andrzej Sicinski: "'Don't Know' Answers in Cross-National Surveys", *Public Opinion Quarterly*, Band 34, 1970, S. 126-129. Die Untersuchung selbst ist dargestellt in Thomas Kutsch: *Die Welt im Jahr 2000*, Scriptor Verlag, Kronberg (Taunus), 1974.
- 169 Siehe hierzu Erwin K. Scheuch: "Society as Context in Cross-Cultural Comparisons", *Social Science Information*, and 6, 1967, S. 7-23.
- 170 Huntington zeigt in einem neuen Bericht einerseits die Fähigkeit der amerikanischen Demokratie, innerhalb von 10 Jahren die Prioritäten der Budgets radikal umzuschichten; während des gleichen Zeitraums ereignete sich ein dramatischer Vertrauensverfall in politische Institutionen, insbesondere in deren Handlungsfähigkeit. Michael Crozier, Samuel P. Huntington und Joji Watanuki: "The Crisis of Democracy", University Press, New York, 1975.
- 171 Vgl. Informationszentrum für sozialwissenschaftliche Forschung: *Forschungsarbeiten 1974 in den Sozialwissenschaften*. W. Kohlhammer, Stuttgart 1975. Inzwischen ist die Zahl der jährlichen Projekte auf über 3000 gestiegen.
- 172 Sozialforschung als Datensammlung, deren Befunde auf maschinell auswertbare Datenträger gespeichert werden, ist dokumentiert in Thomas A. Herz, Hagen Stegemann und

- Sylvia Witte* (Hrsg.): *Empirische Sozialforschung 1974*, Verlag Dokumentation, München 1975. Weniger als 40% der dem Informationszentrum gemeldeten Projekte entsprechen dem vorerwähnten Kriterium. Andererseits ist zu berücksichtigen, daß der gewaltige Umfang kommerzieller Meinungs- und Sozialforschung in diesen Dokumentationen nur in geringem Umfange berücksichtigt werden kann, da die privatwirtschaftlichen Institute nicht sehr auskunftsfreudig sind.
- 173 Zur Konkretisierung dieser Aussage siehe *Erwin K. Scheuch* und *Helmut Zehnpfennig*: "Skalierungsverfahren in der Sozialforschung". In *René König* (Hrsg.), *op. cit.*, Band 3a, 3. Auflage 1974, S. 97-203.
- 174 Die Graphik ist entnommen aus *Herbert B. Asher* und *Bradley Richardson*: *Comparative Voting Behavior*, American Political Science Association, Washington DC, 1975, S. 64.
- 175 Hierüber berichtet *Marcus Felson*: "The Pitfalls of Fancy Analysis – An Exchange in Four Parts", *Sociology and Social Research*, Band 58, 1974, S. 399 ff.
- 176 Eodem loco.
- 177 *Hillel J. Einhorn*: "Alchemy in the Behavioral Sciences", *Public Opinion Quarterly*, Band 36, 1972, S. 367-378.
- 178 *R.M. Guion*: *Personnel Testing*, McGraw-Hill, New York 1965, S. 166.
- 179 *J.S. Armstrong* und *P. Soelberg*: "On the Interpretation of Factor Analysis", *Psychological Bulletin*, Band 70, 1968, S. 361-364.
- 180 *Einhorn*, *op. cit.*, S. 367.
- 181 Die Intensivität gegenüber dem Charakter der Daten wird augenfällig an der Verbreitung der Praxis, die Daten vor der Analyse extrem zu vergrößern, bis hin zur Reduktion auf eine 2 x 2 Tabelle. Damit werden zwar zahlreiche statistische Techniken anwendbar, aber zugleich sind es Techniken, die extrem störanfällig sind. Aus diesen Erwägungen scheint die Haltung der "schwachen Messer" in der Kontroverse über Messniveaus allgemein vorzuziehen; vgl. *Alan C. Acock* und *J. David Martin*, *op. cit.*.
- 182 Ein Beispiel ist *Reinhard Kühnl* u. Mitarb.: *Die NPD. Struktur, Ideologie und Funktion*, Suhrkamp, Frankfurt 1969.