

ZUMA

Zentrum für Umfragen,
Methoden und Analysen

ZUMA-NACHRICHTEN

42

Mai 1998

Herausgeber

Zentrum für Umfragen, Methoden und Analysen (ZUMA)
ZUMA ist Mitglied der Gesellschaft Sozialwissenschaftlicher
Infrastruktureinrichtungen e.V. (GESIS)

Vorsitzender des Trägervereins ZUMA e.V.: Prof. Dr. Max Kaase

Direktor: Prof. Dr. Peter Ph. Mohler

Hausanschrift

B 2, 1
68 159 Mannheim

Postanschrift

Postfach 12 21 55
68 072 Mannheim

Telefon 0621/1246 - 0

Fax 0621/1246 - 100

Redaktion 0621/1246 - 268

E-Mail zuma-nachrichten@zuma-mannheim.de

Internet <http://www.social-science-geis.de/> GESIS
<http://www.zuma-mannheim.de/> ZUMA

Redaktion: Dr. Paul Lüttinger

ISSN 0941-1670 22. Jahrgang

Die ZUMA-Nachrichten informieren über Forschungsarbeiten und Dienstleistungen des Instituts. Namentlich gekennzeichnete Beiträge geben die Meinung der Autoren wieder. Der Nachdruck von Beiträgen ist nach Absprache möglich.

Die ZUMA-Nachrichten erscheinen im Mai und November eines Jahres. Sie werden Interessenten auf Anforderung kostenlos zugesandt.

Druck: Verlag Pfälzische Post GmbH, Neustadt/Weinstraße.
Gedruckt auf chlorfrei gebleichtem Papier.

© ZUMA

In eigener Sache	5
------------------------	---

FORSCHUNGSBERICHTE

Incentives for Survey Participation: Research on Intended and Unintended Consequences. <i>Eleanor Singer</i>	7
Die Messung von Einkommen im nationalen und internationalen Vergleich. <i>Jürgen H.P. Hoffmeyer-Zlotnik und Uwe Warner</i>	30
Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. <i>Achim Koch</i>	66
Analysemöglichkeiten des Mikrozensus. <i>Bernhard Schimpl-Neimanns</i>	91

PROJEKTBERICHTE

„Wienerwald auf Mallorca“: Negative Emotionen von Deutschen gegenüber ihrer eigenen Nation. <i>Thomas Blank und Horst-Alfred Heinrich</i>	120
Zeitbewußtsein, Handlungsintentionen und Eigenverantwortung: Eine Analyse der Zusammenhänge zwischen handlungsbestimmenden Orientierungen. <i>Helga Sievers</i>	144

MITTEILUNGEN

TEXTPACK: Ein Programm zur computerunterstützten Text- und Inhaltsanalyse.....	168
Sozialwissenschaften-Bus 1998/99: Termine und Preise.....	170

BUCHBESPRECHUNGEN

ROBERTS, C. W. (HRSG.): Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts. (<i>Melina Alexa/Cornelia Züll</i>)	174
BÜHL, A./ZÖFEL P.: SPSS für Windows Version 7.5. Praxisorientierte Einführung in die moderne Datenanalyse. (<i>Cornelia Züll</i>).....	178
KIRSCHNER, W.: Episentinel. Monitoring von Krankheiten und Gesundheitsrisiken in Deutschland auf der Grundlage einer Kreistypologie. (<i>Fritz Tiemann</i>).....	179

PUBLIKATIONEN

Buchhinweise	181
ZUMA-Arbeitsberichte	184

ZUMA-TAGUNGEN

Übersicht über die Veranstaltungen im zweiten Halbjahr 1998	187
Workshop: „Einführung in die computerunterstützte Inhaltsanalyse (cui) mit TEXTPACK“, 1.-3. Dezember	187
Konferenz: „Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt“, 8.-9. Oktober	188
Tagung: „Perspektiven einer zukunftsfähigen Ernährungsberichterstattung in Deutschland“, 8.-9. Oktober	193

DURCHWAHL-RUFNUMMERN**ADRESSENPFLEGE****BESTELLUNGEN**

IN EIGENER SACHE

Eleanor Singer vom Institute for Social Research der University of Michigan war im vergangenen Jahr Inhaberin der Rudolf-Wildenmann-Gatsprofessur, die ZUMA zum Gedenken an den Gründungsvorsitzenden von ZUMA seit 1995 jährlich ausschreibt. Sie hielt in diesem Rahmen eine Vorlesung zum Thema „Incentives for Survey Participation: Research on Intended and Unintended Consequences“, die wir in diesem Heft veröffentlichen.

Mit der Problematik der Erfassung von Einkommen bei Umfragen beschäftigen sich *Jürgen H.P. Hoffmeyer-Zlotnik* und *Uwe Warner*. Sie beschreiben unterschiedliche Operationalisierungen in nationalen und internationalen Umfragen und diskutieren Aspekte, die bei vergleichenden Analysen von Einkommen zu berücksichtigen sind.

Unter dem Titel „Wenn ‘mehr’ nicht gleichbedeutend mit ‘besser’ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen“ diskutiert *Achim Koch* die unter Forschern weit verbreitete Annahme, eine Stichprobe sei um so weniger verzerrt, je höher die Ausschöpfung der Umfrage ausfällt. Anhand von sechs allgemeinen Bevölkerungsumfragen, deren Ausschöpfung zwischen 50 und 80 Prozent variiert, untersucht Achim Koch, ob diese Annahme gerechtfertigt ist.

Seit Anfang 1997 können die faktisch anonymisierten Einzeldaten des Mikrozensus 1995 als Scientific Use File vom Statistischen Bundesamt bezogen werden, was mittlerweile viele Forscher nutzen. Neben Vorzügen des Mikrozensus gegenüber sozialwissenschaftlichen Umfragen begegnen die Nutzer auch Einschränkungen, die daraus resultieren, daß das Fragenprogramm vorwiegend durch den Datenbedarf von Verwaltung und Politik bestimmt ist. *Bernhard Schimpl-Neimanns* stellt anhand von Beispielen aus den Bereichen Arbeitsmarkt, Haushalte, Familien und Sozialstruktur vor, wie der Mikrozensus ertragreich ausgewertet werden kann.

Untersuchungen zur nationalen Identität von Personen berücksichtigen selten die Vielfalt der damit verbundenen Gefühle. So werden z.B. negative Emotionen zwar in theoretischen Überlegungen thematisiert, empirisch aber mangels geeigneter Instrumente bislang nicht systematisch untersucht. *Thomas Blank* und *Horst-Alfred Heinrich* stellen die Entwicklung eines Meßinstrumentes vor, welches das Peinlichkeitsgefühl gegenüber der eigenen Nation erfaßt.

Helga Sievers behandelt in dem Beitrag „Zeitbewußtsein, Handlungsintentionen und Eigenverantwortung“ das Problem, daß Zeit von den einzelnen Menschen auf unterschiedliche Art wahrgenommen und interpretiert wird und welche unterschiedlichen Implikationen sich hieraus für ihre Handlungsintentionen ergeben.

Bei den Personalien haben sich folgende Änderungen ergeben: Neue Mitarbeiterinnen bei ZUMA sind *Patricia Lüder* (Abteilung Textanalyse, Medienanalyse und Vercodung) und *Sonja Hess* (Abteilung Soziale Indikatoren). Ausgeschieden sind *Matthias Engelhard*, *Ursula Palm* und *Ingrid Weickel*.

Ich hoffe, daß unser Aufsatz- und Informationsangebot das Interesse der Leser weckt und wünsche viel Vergnügen und Gewinn bei der Lektüre.

Max Kaase

Vorsitzender des ZUMA e.V.

INCENTIVES FOR SURVEY PARTICIPATION: RESEARCH ON INTENDED AND UNINTENDED CONSEQUENCES^{*)}

SECOND ANNUAL WILDENMANN LECTURE

Eleanor Singer

I am honored to have been invited to give the second Rudolph Wildenmann lecture, following in the distinguished footsteps of Norman Bradburn last year. Although the topic of my talk is perhaps a little specialized, it is in the spirit of empirical social research, especially survey research, which Professor Wildenmann did so much to foster.

As you all know, sample surveys are increasingly used by government and the private sector to inform policy decisions in a wide variety of areas. They are also indispensable for theory building and theory testing in the social sciences.

But while reliance on the products of survey research is growing, the difficulty of obtaining cooperation with sample households in the United States and other developed countries is increasing as well (de Heer and Israëls, 1992). Even though the overall response rates of surveys may be relatively constant over time (Smith, 1995), the component of nonresponse due to refusals appears to be increasing unless efforts by survey organizations are substantially increased (Groves and Couper, 1996).

One of the strategies increasingly being used by survey organizations is the payment of monetary incentives. But payment of such incentives, either at the outset of the survey or in an effort to convert refusals later, raises a variety of long-term issues for survey organizations.

Shortly before I came to the Institute for Social Research at the University of Michigan in the fall of 1994, Professor James House, the Director of the Survey Research Center, asked Bob Groves and me to organize a research seminar that would discuss some of these issues. Following presentations by four distinguished speakers, including Barbara Bryant, a former director of the U.S. Census Bureau, and Tom Juster, an economist and

Senior Research Scientist at ISR, several of us met in an effort to formulate guidelines for an SRC policy on the payment of incentives to respondents. One of the few things the panel agreed on was the need for continued research, both on the effects of incentives and on public perceptions of, and reactions to, them.

Today's talk is about what I think we have learned since then about the role of incentives in motivating survey participation, at least in the United States. It remains to be seen whether the findings will generalize to other countries, with different traditions and different social norms. My hope is that this talk will stimulate a similar program of research at ZUMA, and elsewhere.

I apologize that this talk is in English rather than in German. But I am going to try to tell you a story with pictures, and those of you who are interested can read the lecture at your leisure. The lecture has three parts:

1. First, it reviews what is known about the effect of incentives on response rates in mail surveys, and describes recent research on how they affect face-to-face and telephone surveys.
2. Second, it discusses issues of equity that arise in connection with refusal conversion payments, and reviews recent research on how respondents perceive such issues and whether or not these perceptions affect their willingness to participate.
3. Third, it describes some empirical evidence on whether or not the payment of monetary incentives is having long-term effects on the motivation for survey participation.

Another way of stating this is to say that the lecture deals with *intended and unintended* consequences of paying incentives to respondents.

From the outset, I want to make it clear that I will be talking about a program of research involving the collaboration of many people, in particular my colleagues Bob Groves, Trivellore Raghunathan, John Van Hoewyk, and my research assistant, Amy Corning. The National Science Foundation and the Survey Research Center have provided indispensable financial support.

What Do We Know about How Incentives Work?

Effects in Mail Surveys

From a meta analysis of the effect of incentives on mail survey response rates (Church, 1993) we know that:

- 1) Prepaid incentives yield higher response rates than promised incentives;
- 2) Prepaid monetary incentives yield higher response rates than gifts;

- 3) Response rates increase with increasing amounts of money, although not all studies find a linear relationship;
- 4) Incentives promised after return of the questionnaire do not significantly increase response.

Effects in Telephone and Face-to-Face Surveys¹⁾

Initial incentives--that is, incentives offered to all respondents at the outset of a survey--are also being used increasingly in telephone and face-to-face surveys, but we know much less about their effectiveness in these modes. Comparison of the respondent's situation in mail and interviewer-mediated surveys suggests that, other things being equal, the need for a monetary incentive should be greater in mail surveys. This conclusion is based on the persuasive value generally attributed to the interviewer, as well as on the fact that the presence of an interviewer lessens the respondent's burden in completing the questionnaire. Thus, it seemed possible that the significantly positive effects of incentives documented for mail surveys would not be found in the case of interviewer-mediated surveys.

Nancy Gebler, John Van Hoewyk, Trivellore Raghunathan, Kate McGonagle and I brought together and quantitatively analyzed all the experimental studies we could locate (as of 1997) which used incentives in face-to-face and telephone surveys--37 studies in all--each containing two, and sometimes more, experimental conditions (Singer et al., forthcoming). We attempted to answer four basic questions:

- Do incentives improve response rates in telephone and face-to-face surveys? By how much?
- Are prepaid incentives more effective than promised incentives?
- Is money more effective than a gift?
- What is the effect of burden?

For each experimental condition in each of the 37 studies, the following independent variables were coded:

- Amount of incentive
- Type of incentive (gift or money)
- Timing of incentive (paid before interview taken, or promised)
- Burden (coded as high if interview was longer than an hour or if any of the following were present: diary, test, sensitive questions, panel study; otherwise coded as low)

The dependent variable was the difference in response rates between the zero incentive condition (or the lowest incentive condition) and each higher incentive condition in the

same experiment. So, for example, if a study had three experimental conditions--a zero-incentive condition, a \$5 incentive, and a \$10 incentive--we computed two dependent variables for that experiment: the difference in response rate between the zero-incentive and the \$5 condition, and the difference between the zero-incentive and the \$10 condition. These differences were weighted by the inverse of the variance of the response rate difference in the analysis.

In order to eliminate confounding due to differential cooperation tendencies, we introduced a covariate which we call "panel status" into all the analyses reported below. Panel status simply refers to whether the subjects in the experiment were new respondents, nonrespondents to the current or a prior wave, or respondents to a previous wave. Most of the experiments were carried out with respondents who had not previously been approached for an interview. Because these variables were not significant in any of the models, we did not include them in the final models shown. Similarly, we included mode of administration as a control variable, but trimmed it from the models when it failed to show significant effects.

The results of the analysis are shown in Table 1. The first model shown in Table 1 compares one or more incentive conditions with a zero incentive condition in the same experiment. The model controls for type of incentive, timing of incentive, amount of incentive, and burden. As can be seen from Model 1, paying an incentive has a significant positive effect with these other variables controlled. As in the study by Yu and Cooper (1983) and Church (1993), the effects of incentives are linear; a term for the square of the incentive difference, introduced to test for curvilinearity, was not significant. Within the limits of incentives and response rates occurring in these experiments, more money results in higher response rates.

The difference in response rates produced by prepayment vs. promised payment is not significant in this model, but the direction of the differences is in accord with conventional wisdom--i.e., prepayment appears to be more effective than promised payments. When we look only at the five experiments that hold size of incentive constant and compare prepayment and promised payment directly, we find in every case that prepayment yields higher response rates than the promised payment condition. However, unlike Church (1993) but like Yu and Cooper (1983), we found a significant improvement in response rates with promised as well as prepaid incentives (data not shown).

Like Church (1993), we found that gifts are significantly less effective in eliciting response, even controlling for the value of the incentive. Increasing the burden of the interview significantly increased the difference in response rates between a zero

incentive and an incentive condition. The interaction between burden and the size of the incentive was not significant, however.

Are incentives necessary in low-burden telephone or face-to-face interviews? We cannot answer the question in that form, but we can ask whether incentives have a significant effect on response rates in such interviews. Restricting the analysis above to low-burden conditions, we find that the effect of incentives remains significant ($b=0.5865$, $S.E=0.2144$, $p<.05$).

The second model in Table 1 includes, in addition to experiments with a zero-incentive condition, all those that compare the smallest with larger incentive conditions as well. These results parallel those shown in Model 1, except that here the effect of prepayment is not only positive but also significant. The effect of gifts is significantly negative (compared with cash), and the effects of burden are as described for Model 1.

Table 1: Effects of Incentive Amount, Time of Payment, Type of Incentive and Burden on Differences in Response Rate

	Model 1		Model 2	
	<i>Zero vs. Larger Incentive Conditions</i>		<i>Smallest vs. Larger Incentive Conditions</i>	
	<i>Beta</i>	<i>Standard Error</i>	<i>Beta</i>	<i>Standard Error</i>
Intercept	1.7852	1.2808	1.4492	1.2785
Incentive difference/dollar ^a	0.1479	0.0479***	0.1443	0.0481***
Prepayment	0.4599	2.0293	2.4462	1.5066*
Gift	-4.9357	2.1360**	-7.1428	1.4257***
Burden	3.2866	1.3308**	3.9550	1.2289***
Adj. R ²	.36		.41	
N	(55)		(59)	

^a „Incentive difference“ refers to the difference between the smallest (or zero) incentive and larger incentive amounts paid. Incentive amounts are modeled in constant 1983 dollars.

* $p = .11$, ** $p < .05$, *** $p < .01$

Because some studies contributed more than one pair of conditions to the analysis, we reran these analyses using a random effects model with studies as random blocks. The parameter estimates are very similar to those in models that do not control for study, and the effects of all the independent variables except prepayment remain significant in these models.

As far as we can tell from the studies we examined, incentives do not appear to exact a price in quality (i.e., item nonresponse or number of words in response to open-ended

questions). Paying an incentive may alter the composition of the sample, however. In four studies, there is an indication that paying an incentive may be useful in obtaining higher numbers of respondents in demographic categories that might otherwise tend to be underrepresented in sample surveys (e.g., low income or nonwhite race). Six other studies reported no significant effects of incentives on sample composition, and in one study the results were mixed.

Does \$10 Equal \$10? The Effects of Framing on the Effectiveness of Incentives

The theoretical literature on the norm of reciprocity (Gouldner, 1960; Cialdini, 1988) suggests that framing an appeal in such a way as to evoke a sense of diffuse obligation will result in higher response rates than framing it as payment for the respondent's time. The difference can, perhaps, also be thought of as a *Gemeinschaft* vs. *Gesellschaft* orientation (Toennies, 1955). The former is a reflection of a community of interests; the latter fosters a rational calculation of benefits and costs.

The question we addressed in the study I am about to describe is whether this theoretical principle depends on the type of incentive offered (Singer, Gebler, Van Hoewyk, and Brown, 1997). That is, we hypothesize that the use of a monetary incentive renders an appeal to reciprocity less effective, and that the use of a gift renders it more effective. In other words, we hypothesize that certain types of appeals are congruent only with certain types of incentives, and that they are less effective when this natural congruity is violated.

In order to test this hypothesis, we designed an experiment in which two types of incentives (a pen valued at \$3.75, and a check for \$10) were randomly paired with two different types of appeals.

One letter described the incentive as a "token of our appreciation"; the other referred to the incentive as "payment for your time." A random half of the group receiving each type of incentive received the "token" letter, the other half received the "payment" letter. Aside from the definition of the incentive as a token of appreciation or as payment, the two letters were identical in describing the content of the study (how to make surveys more effective), its sponsor (the University), and the organization carrying it out (the Survey Research Center at the University of Michigan). We predicted that describing a pen as a "token of appreciation for your help" would be perceived as an appeal to the norm of reciprocity, and would be more effective in eliciting agreement to participate in the study than a pen described as "payment for your time." Similarly, we predicted that a check for \$10 described as "payment for your time" would be more effective in eliciting agreement to participate than a check for \$10 framed as "a token of our appreciation."

Since most monetary incentives in face-to-face and telephone surveys are framed as promised rather than prepaid incentives, we also included a condition in which the \$10 check was described as a payment but promised to the respondent rather than sent ahead of time, so that we would be able to distinguish the effects of prepayment from the effects of appropriate framing. 125 undergraduates from the University of Michigan were randomly assigned to each of the five conditions.

Students were mailed a letter asking them to participate in a telephone survey being conducted by the Survey Research Center of the University of Michigan about how to improve survey practices. The survey was described as taking 45 minutes, and students were asked to return a card indicating their willingness to be interviewed. All students (both those who returned the card and those who did not) were followed up with a brief telephone interview. The questionnaires differed slightly between the two groups, but both contained similar sets of questions about their attitudes toward surveys and incentives and their reasons for returning or failing to return the card.

Looking only at those who returned a card indicating their willingness to be interviewed supported our framing hypothesis for money, but not for gifts (see Table 2, Column 1).²⁾ As in other reviews of the role of incentives in mail as well as telephone and face-to-face surveys (Church, 1993; Singer et al., forthcoming; Yu and Cooper, 1983), we found that money is significantly more effective than a gift (57.6% vs. 32.0%; $X^2=32.1$, $df=1$, $p<.01$), and promised payment is significantly less effective than a prepaid monetary incentive (44.8% vs. 57.6%; $X^2=8.52$, $df=1$, $p<.01$).³⁾

As predicted, framing a check as payment is significantly more effective than framing it as a token of appreciation (64.0% vs. 51.2%; $X^2=6.54$, $df=1$, $p<.05$), but there is no "boomerang" effect when a pen is framed as "payment" for time--perhaps because the incentive effect associated with the pen is so small.

As already noted, this conclusion is based on those expressing a willingness to participate by returning a card. In fact, however, we attempted to interview everyone in the sample, regardless of whether or not they had mailed back the card. Response rate to the survey can be seen as a measure of how firm the decision to participate or refuse, represented by mailing back the card, actually is.

Among those who had returned a card signifying willingness to participate, the response rate to the actual survey was 92%, indicating that mailing the card constitutes a high, but not perfect, commitment to complete the interview. But even among those who had not returned the card, 69% were ultimately interviewed.⁴⁾ Thus, the negative commitment signified by failure to return the card is not very firm; persuasive interviewers can over-

come it in a large majority of potential respondents. We should note once again that these findings are based on a student sample, notoriously more cooperative than a general population sample, and that everyone in the sample either had received (80%) or was promised (20%) some kind of incentive.

The actual survey response rates, by experimental condition, are shown in Table 2, Column 2. Note that these cannot be taken as a pure measure of incentive effectiveness, since the incentive is, in every case, augmented by one or more contacts with an interviewer. Thus, Column 2 illustrates the additive and interactive effects of interviewers and different kinds of incentives. In the case of some of the experimental conditions, this "added value" is very large indeed: more than forty percentage points in the case of the pen incentive conditions and almost forty percentage points in the promised money condition.⁵⁾

Table 2: Agreement to Be Interviewed and Actual Response Rates, by Experimental Condition

<i>Experimental Condition</i>	(1)	(2)	<i>(N)</i>
	<i>Agreement to Be Interviewed</i>	<i>Response Rate</i>	
	%	%	<i>N</i>
Pen as a token of appreciation	30.4	76.0	125
Pen as payment for time	33.6	75.2	125
\$10 check as token of appreciation	51.2	79.2	125
\$10 check as payment for time	64.0	81.6	125
\$10 check as payment, promised	44.8	84.0	125

However, it is not simply true that the response rates after follow-up are higher than those without interviewer intervention. *The conclusions one would draw about the relative effectiveness of different incentives also differ.* None of the differences between conditions is significant after the calls by interviewers. Not only is there no support for the framing hypothesis when only the results of the actual survey are considered, but there is no support for the effects of cash vs. gifts or prepaid vs. promised incentives.

Because these results are based on an experiment with students, and because of several other ambiguities, including the absence of a control group, we are in the process of repeating this experiment with a national sample of telephone households for whom we were able to obtain a listed address. So far, the results resemble the results for the student sample before interviewer followup, except that the response rates, as one would expect, are significantly lower.

Does Equity Matter? And to Whom?

So far, we have been looking at research on intended consequences of incentives. I turn now to some experiments involving *unintended* consequences.

There are two opposing views on the issue of paying refusal conversion payments to reluctant respondents. On one side are the economists, naive or otherwise, who argue that refusal is an *ipso facto* indication that the survey has less utility for the refusers, or is perceived as more burdensome by them. This being the case, it is seen as entirely appropriate to offer compensation to refusers but not to other respondents, who are more cooperative.

On the other side of the argument are the social psychologists, for whom refusal conversion payments represent a violation of equity expectations. For the social psychologists, reluctance to participate is not an *ipso facto* indication that the survey is more burdensome; and offering refusal conversion payments to reluctant respondents seems likely to reinforce their uncooperative behavior, making it more likely to recur in the future. In addition, such payments are believed likely to alienate cooperative respondents if they should become aware of the practice.

In the traditional formulations of equity theory (e.g. Adams, 1965; Adams and Freedman, 1976; Walster, Berscheid and Walster, 1976; Walster, Walster, and Berscheid, 1978; Homans, 1961, 1974), perceptions of equity are determined by the results of a comparison such as the following:

$$\frac{\text{Ego's Input}}{\text{Ego's Reward}} = \frac{\text{Alter's Input}}{\text{Alter's Reward}}$$

If the two ratios are similar, the outcome of the comparison is perceived as equitable. But if Alter is seen as differentially rewarded for a task similar to that performed by Ego, then the situation is likely to be perceived as inequitable, *provided that the other person is seen as "similar" to Ego*. Thus, offering incentives to respondents who have refused a prior survey request may be perceived as inequitable by those who have agreed to participate without such an incentive. Others, however, may not be defined as similar, or the task they perform may be seen as dissimilar. For example, potential respondents may recognize that although *they* are motivated to respond by altruistic concerns, others are not so motivated, and it is therefore appropriate to provide greater incentives to those others. Or, potential respondents may judge that the task is harder for reluctant respondents, and that it is therefore appropriate to offer greater incentives to them. Under those circumstances, if other respondents are not seen as similar, then reference

group theory (Hyman, 1942; Stouffer et al., 1949; Merton and Rossi, 1950; Hyman and Singer, 1968), cognitive dissonance theory (Festinger, 1957), and social comparison theory (Suls and Miller, 1977) all predict that they will cease to be defined as an appropriate object of comparison, and differences between Ego's reward and Alter's will cease to motivate Ego's behavior (Suls and Miller, 1977; Singer, 1981; Shepelak and Alwin, 1986).

If an expectation exists concerning appropriate rewards for survey participation, then it is reasonable to assume that a violation of that expectation by the survey organization will lead to a perception of inequitable treatment by the respondent. In turn, such a perception may lead to reluctant participation or outright refusal.⁶⁾

Focus Groups

To explore potential respondents' reactions to equity issues involved in the payment of incentives, we held two exploratory focus group discussions in the fall of 1994. One consisted of black low-income participants, all but two women, in Detroit; the other, of middle-class, mostly white, men and women in Livonia, a suburb of Detroit. Respondents were recruited by Survey Research Center interviewing supervisors and promised \$15 for their participation.

These group discussions offered little evidence for widespread perceptions of inequity in survey organizations' incentive practices. But to provide better answers to this question, Bob Groves and I designed a study that relied in part on laboratory administration of videotaped vignettes of survey interactions and in part on a field experiment in the 1996 Detroit Area Study (Groves et al., 1997).

The Videotaped Vignette Study

The vignettes were constructed from three completely crossed factors: disclosure of differential incentives (vs. no disclosure), burden (high vs. low), and traditional persuasion (no prior refusal vs. prior refusal). In addition, we included two other factors that were only partially balanced within the design: a rationale for differential incentives in the disclosure condition (vs. no rationale), and an initial incentive (vs. no incentive). The result was 12 distinct vignettes, having the characteristics shown in Figure 1. These were videotaped, using one interviewer throughout, a second interviewer when a refusal conversion was involved, a female subject's voice, and a male subject's voice. Ninety-six subjects were recruited through the Parent-Teacher Organizations of two local elementary schools to which we made a cash contribution; subjects were not paid

individually. They were interviewed a week or so before viewing the videos with a questionnaire that asked about their political attitudes and community involvement, as well as about their attitudes toward surveys and their sensitivity to equity issues. Each subject viewed a randomly selected set of six vignettes and completed a brief self-administered questionnaire after seeing each one. The questionnaires included some manipulation checks along with the main dependent variable: "Would you be willing to take part in this interview?" Following the last video, they also completed a "final debriefing questionnaire" which repeated a few of the questions from the intake interview, primarily those about their perception of the fairness of differential payments to survey respondents.

Figure 1: Subject Counts by Experimental Factors

	No Incentive			
	Low Burden		High Burden	
	No Prior Refusal	Prior Refusal	No Prior Refusal	Prior Refusal
No Disclosure	48	48	48	48
Disclosure, No Rationale	48	48	48	48
Disclosure, Rationale	48	X	48	X
	Incentive			
	Low Burden		High Burden	
	No Prior Refusal	Prior Refusal	No Prior Refusal	Prior Refusal
No Disclosure	48	X	48	X
Disclosure, No Rationale	X	X	X	X
Disclosure, Rationale	X	X	X	X

The analysis controls for order-of-administration effects (not significant) as well as mode effects (25 subjects responded to written, rather than videotaped, versions of the vignettes). Although we also entered demographic and attitudinal variables into the logistic regressions, only one of these produced a significant effect, either alone or in interaction with the experimental variables, and is retained in the final model. The results of the final analysis, which was run in SUDAAN to adjust for multiple observations on the same respondent, are shown in Table 3.

The table shows some obvious things and some that are less obvious:

1. In the first place, interviews described as requiring an hour and a half are significantly less likely to elicit a willingness to respond than those described as requiring only half an hour.

2. Interviews described as offering an incentive ahead of time significantly increase the expressed willingness to respond.
3. Showing a vignette in which a respondent refuses twice, and is solicited three times by two different interviewers, *reduces* the expressed willingness to respond, as we had hypothesized; the effect is significant at the zero-order level at $p=.09$.
4. Subjects answering the written, rather than the videotaped, versions of the vignettes are *less* likely to express their willingness to respond, perhaps because of the absence of social desirability or identification factors, but there are no significant mode by independent variable interactions.
5. Subjects who believe respondents should be paid for survey participation were less likely to express a willingness to respond.
6. Finally, and most important from the point of view of equity theory, disclosure of differential incentives being paid to some but not all respondents significantly reduces subjects' expressed willingness to respond. However, if the vignette contains a rationale for why differential payments are necessary, the negative effect of disclosure is much reduced.

Table 3: Logistic Regression Coefficients for Model Predicting High Likelihood of Participation ¹⁾

	Coefficient	Standard Error
Intercept	.11	.28
Predictors		
Incentive	.94**	.27
Burden	-.58*	.26
Disclosure, Rationale Conditions:		
No Disclosure ²⁾	---	
Disclosure, No Rationale	-.59**	.18
Disclosure, Rationale	-.25	.23
Refusal	-.49	.32
Mode	-.75*	.37
Compensation	-.85*	.40

1) Standard errors reflect the clustering of observations within subjects. The dependent variable is coded 0 (for responses 1, 2, 3) and 1 (for responses 4, 5). The *compensation* variable is coded "1" for "strongly agree" or "somewhat agree" and "0", otherwise. Asterisks denote p-values: * - $p<.05$, ** - $p<.01$.

2) Omitted category in dummy variable regression.

Not shown in Table 3, but of interest nevertheless, is the fact that there is no independent or interactive contribution of a measure of perceived fairness on expressed willingness to participate. That is, although about 80% of subjects consider the paying of differential incentives unfair (unless justified by greater importance of the information or a longer interview), this attitudinal measure does not increase our ability to predict their willingness to participate in a survey in which differential incentives are being paid. Furthermore, the percentage of those considering the practice unfair declines somewhat from intake to debriefing interview.

The Field Experiment

Because the vignette study was carried out in the laboratory with a rather homogeneous sample (most respondents were white married women with young children, members of two Parent-Teacher Organizations in Ann Arbor who had been recruited through the promise of a contribution of \$200 to the PTO), we designed a field experiment that, we hoped, would provide better external validity for these findings.

The vehicle for this investigation of the role of perceptions of equity in survey response was the Detroit Area Study (DAS), established by the University of Michigan over 40 years ago as a research and training laboratory in the community. The 1996 DAS was intended as a study of interracial contacts and attitudes, but the survey was also designed to permit an investigation of hypotheses concerning perceptions of equity in relation to the payment of incentives.

The 1996 DAS consisted of an area probability sample of 451 households, interviewed in person between April and August 1996 by graduate students supplemented, during the summer months, by professional interviewers from the Survey Research Center at the University of Michigan. The completion rate for the study, defined as the proportion of interviews completed divided by the number of households listed less vacants and non-sample households and those ineligible because of illness or language problems, was 66%. Seventy-two percent of the respondents were white, 21% were black, and 7% some other race. Ages ranged from 18-89 with a mean age of 44 years. Respondents had a mean of 13.5 years of education and the median income fell in the \$40,000-\$44,999 range.

All households were sent an advance letter before the interviewer contacted them; two thirds of these letters contained a \$5 bill and one third did not.⁷⁾ Following one refusal by anyone in the household, these two portions of the sample were again split in half. Half of the refusers in each group--those who had received an initial incentive, and those who had not--were sent a persuasion letter only; the other half were offered \$25 to

convert their refusal. Thus, ultimately, there were four groups of respondents, of unequal size: Those who had been offered nothing and who had given an interview; those who had given an interview after being offered \$5; those who had given an interview after being offered \$25, and those who consented after receiving \$5 and being offered \$25.

Near the end of the interview, we introduced a split-ballot experiment into the survey. After a few questions about the present survey, half the subjects in each of the four experimental groups above were informed that because of the importance of including everyone in the sample, some of those who had expressed reluctance to participate had been offered \$25, while others received nothing; and they were asked whether they considered this practice fair or unfair. For the other half of the sample, this piece of information, and the question about fairness, were omitted. Then, all subjects were asked about their willingness to take part in a future interview by the same organization.

We had expected those to whom the payment of differential incentives was disclosed to be less willing to participate in a future survey by the same organization, because they would consider the practice unfair. However, at the zero-order level, there were no differences in expressed willingness to participate between the group to which differential payments had been disclosed and the group to which it had not--about a quarter of each group said they would "definitely" be willing to participate in another survey by the same organization, even though a large majority--72.7%--of respondents to whom the payment of differential incentives was disclosed considered the practice to be unfair.

We had also hypothesized that respondents to whom differential incentive payments were disclosed *and who perceived these payments to be unfair* would be less willing to participate in future surveys than those who perceived them as fair. Although a somewhat greater percentage of those who considered the practice fair than of those who considered it unfair were "definitely" willing to participate in a future survey by the same organization (32.8% vs. 25.8%), this difference was not statistically significant ($X^2=1.08$, $df=1$, $p<.30$).

Because it seemed possible that some of the variables specified above might interact, we estimated several logistic regression equations from the data controlling for these other variables as well as a series of demographic characteristics. None of them altered the basic conclusions drawn from the bivariate relationships.

A little more than a year later, all the original DAS respondents for whom we could locate an address were sent a mailed questionnaire on the topic of assisted suicide. The return address on the letter differed from that used for the original Detroit Area Study survey, complicating the inferences that can be drawn from the experiment about the

effect of concerns about equity.⁸⁾ The questionnaire was preceded by an advance letter; a random half of the respondents received a \$5 prepaid incentive with the letter, whereas the other half did not. A week after the questionnaires had been mailed, everyone in the sample was sent a reminder postcard. We did not engage in further efforts to stimulate the response rate.⁹⁾

The response rate to the self-administered survey was 41.4% for those who did not receive an incentive and 65.1% for those who did--a large and statistically significant difference. Our real interest in the study, however, was the consequence of disclosing differential payments to respondents for their survey participation a year later. Such disclosure appears to have some effect on reducing participation a year later: 50% of those to whom differential payments were disclosed participated in the new, self-administered survey, compared with 55% to whom such payments were not disclosed, but this difference, while in the expected direction, is not statistically significant. When we looked at the relationship between disclosure and participation in a logistic regression framework, adding controls for receipt of an incentive in the DAS, receipt of a refusal conversion payment offer, and a series of demographic variables, the effect of disclosure remained constant but did not reach a conventional level of statistical significance ($b=-.318, s.e.=.23, p=.17$).

Are Incentives Transforming the Motivation for Survey Participation?¹⁰⁾

I come now to the last topic of the lecture: Are incentives transforming the motivation to participate in surveys?

Because of concerns about declining cooperation rates, the Survey Research Center at the University of Michigan decided in the fall of 1995 to begin monitoring the changing climate for survey research in the United States by adding five evaluative questions at the end of the Survey of Consumer Attitudes, a national telephone survey administered monthly to a sample of roughly 500 respondents. Of these 500, 300 are newly selected RDD households and the remaining 200 are reinterviews of respondents first interviewed six months earlier. Because of concerns about their possible biasing effects, the five evaluative questions were asked of only the reinterviewed portion of the sample. The questions were added to the survey in January and February of 1996 and repeated in February and March of 1997 in order to measure changes in the climate for survey research over the approximately 12-month period. In principle, one could use changes in the responses to these questions as leading indicators of changes in the climate for survey research, and take proactive steps to counteract such changes.

Response Rate

The analysis involves comparing two independent samples of reinterviewed respondents. The initial response rate of those reinterviewed in January and February of 1996 was 70.3% six months earlier; their reinterview rate averaged 77.6%. Thus, the effective response rate of the 1996 sample is 54.5%. For the sample reinterviewed in February and March of 1997, the initial response rate averaged 69.3% and the reinterview rate, 76.8%; thus, the effective response rate for the 1997 sample was a slightly lower 53.2%. (The response rate excludes only nonsample cases from the denominator, and is thus a fairly conservative estimate. Noninterviews for reasons of illness or language, for example, are retained in the denominator.) The questions are, thus, asked primarily of cooperative respondents; presumably, the attitudes toward surveys among nonrespondents would be more negative.

Monitoring Questions

The five monitoring questions asked on the surveys are as follows:

1. If you had it to do over again, would you have agreed to do the interview or would you have refused?
2. For each of the following, please tell me whether you agree strongly, agree somewhat, disagree somewhat, or disagree strongly. Surveys like this one provide useful information for decision makers?
3. Surveys like this one are a waste of people's time.
4. People should get paid for doing surveys like this.
- 4a. How much should they get paid?
5. Everyone has a responsibility for answering surveys like this. Do you agree strongly, agree somewhat, etc.

The five questions were systematically rotated during their administration.

Changes over Time

As already noted, the reason for asking these questions at two points in time was to capture changes in the climate for surveys. Responses to the five questions (and one subquestion) in 1996 and 1997 are shown in Table 4; they represent two cross-sectional measurements of attitudes toward surveys rather than answers by the same respondents at two different times.

Table 4 indicates that on three of the questions, no significant changes took place from one year to the next. Three others, however, show a significant change: Significantly

more respondents (45.7% in 1997, compared to 29.7% in 1996) said that respondents should be paid for doing a survey like this, and the amount they stipulated showed a significant increase as well. At the same time, significantly more respondents (51.4% in 1997, compared with 44.5% in 1996) agreed that everyone had a responsibility to answer a survey like this.

We had anticipated changes in answers to the question about payment for two reasons. First, the practice of paying incentives to respondents in telephone and face-to-face surveys appears to be increasing, and this may be influencing expectations among the general public. Second, a large number of respondents to the March 1997 survey had themselves received an initial incentive payment of \$5 six months earlier, and a smaller number of respondents in three of the four months had received \$20 or \$25 to convert an initial refusal.

Table 4: Responses to Evaluative Questions, by Year

	1996 (%)	1997 (%)
R Should Get Paid **		
Agree	29.7	45.7
Disagree	70.3	54.3
(n)	(411)	(396)
How Much *		
0-5	19.6	22.2
6-10	32.4	19.4
11-20	19.6	35.4
Over 20	28.4	22.9
(n)	(102)	(144)
Surveys Are Useful		
Agree Strongly	34.6	39.7
All Other	65.4	60.3
(n)	(405)	(401)
Everyone's Responsibility *		
Yes	44.5	51.4
No	55.5	48.6
(n)	(409)	(403)
Will Do Again		
Yes	76.9	76.1
No	23.1	23.9
(n)	(407)	(406)
Waste of Time		
Disagree Strongly	28.6	30.4

All Other (n)	71.4 (402)	69.6 (395)
------------------	---------------	---------------

* p <.05, **p <.01

In order to separate the effect of generalized expectations from that of personal experience, we looked at 1997 responses to the "People should get paid" question among those who had and those who had not been offered an incentive. The results are shown in Table 5. In both years, those who had received an incentive were much more likely to say that people should be paid than those who had not; the differences are significant in both years, and are especially large for people who received the (larger) refusal conversion payments. Among people who did *not* receive any kind of incentive in either year, there is a slight increase from 1996 to 1997 in the percentage saying people should get paid for doing surveys like this, but this difference is not significant; $X^2=1.77$, $df=1$, $p=.18$. Nor is the difference between 1996 and 1997 among people who received a refusal conversion payment significant. Thus, Table 5 demonstrates that the changed expectations apparent in Table 4 are due almost entirely to the responses of those who had themselves received an initial incentive--in other words, to personal experience rather than diffuse social norms.

The question of interest for this paper, however, is whether these responses should be understood as *reflecting changed expectations for the future*, or as *normative statements justifying past behavior*.

Table 5: Responses to Evaluative Questions, by Incentive and Year

	1996			1997			
	Incentive			Incentive			
	0	\$20-25		0	5\$	\$20-25	
R Should Get Paid	26.0%	91.3%	***	31.0%	51.3%	77.6%	***
Surveys Are Useful	32.9%	63.6%	***	35.2%	48.7%	40.6%	
Everyone's Responsibility	44.8%	39.1%		45.7%	59.0%	58.0%	*
Will Do Again	77.6%	65.2%		74.5%	83.3%	76.7%	
Waste of Time (Disagree Strongly)	26.6%	63.6%	***	30.0%	38.5%	26.5%	
N	(393)	(23)		(242)	(81)	(70)	

* p <.10, ** p <.05, *** p <.01

We can test these alternative interpretations by examining the cooperation rate of people to the March 1997 survey. Among people recontacted in March 1997, 139 had received an incentive 6 months earlier and 98 had not. If the earlier payment of an incentive led

to (unmet) expectations for payment in the future, cooperation rates (without an additional incentive) in March should be lower among those who had received an initial incentive the preceding September than among those who had not. However, among those who had received an initial incentive in September and who were contacted by interviewers, 81.0% were reinterviewed without an additional incentive in March; among those who had received a refusal conversion incentive in September, the cooperation rate in March without an additional incentive was 79.5%; and among those who had received no incentive in September, the cooperation rate without an additional incentive in March was 66.3%. The difference between those receiving no incentive in September and those receiving \$5 is significant; $X^2 = 5.43$, $df=1$, $p<.05$; those who had received a five dollar incentive six months earlier were significantly *more* likely to cooperate in March than those who had received no incentive. The difference in March cooperation between the no-incentive group and the group receiving a refusal conversion payment in September is not significant. Thus, these data provide no support for the interpretation that responses to the question about whether or not respondents should get paid reflect expectations about future behavior on the part of the survey organization.

Summary and Conclusions

What do these investigations tell us about the anticipated and unanticipated consequences of incentives?

1. As far as anticipated consequences are concerned, incentives appear to have similar effects in interviewer-mediated surveys as they do in mail surveys. They increase response rates, even for low-burden surveys. They do not appear to have adverse effects on the quality of response, although they may alter the composition of the sample. Conclusions about the effects of framing await further testing, but results so far suggest they are relatively unimportant.
2. As far as unintended consequences are concerned, the evidence is more ambiguous. Our research suggests that equity matters to respondents, but it is not a very salient issue for them. Thus, although large majorities consider the practice of paying refusal conversion payments unfair, awareness of this practice does not seem to loom large in their expressed willingness to participate in future research by the same organization or in their actual participation a year later. Nor does the payment of incentives at the outset of a survey appear to have negative consequences for participation in a second wave of the same survey without additional incentives. On the contrary, people who received such an incentive six months earlier were actually more likely to cooperate than those who did not. However, payment of incentives is still a rather novel experience for respondents to

telephone or personal interviews. Whether the absence of negative results observed in the present study will survive the more widespread use of incentives in such interviews remains very much an open question, and one deserving continued research.

3. The general conclusion, therefore, should come as no surprise. Although we have not so far seen evidence that the payment of incentives, either to some or to all respondents, has dramatic unintended consequences, we cannot rule out the possibility of such effects in the future. Nor can we be sure that the effects would be the same in countries with different cultural norms about equitable treatment and payment for survey participation. Therefore, experiments along the lines reported here should be carried out in other countries, and we should continue to monitor attitudes toward surveys both in the U.S. and in other countries where surveys are routinely carried out and increasingly relied on by governments, businesses, and academics.

Author's Address

Dr. Eleanor Singer
Survey Research Center
Institute for Survey Research
Box 1248
Ann Arbor, MI 48106
email: esinger@isr.umich.edu

Notes

*) Anmerkung des Herausgebers: Der Beitrag ist die überarbeitete Fassung einer Vorlesung, die Eleanor Singer im Rahmen der Wildenmann-Gastprofessur bei ZUMA hielt.

1) A fuller report of the research in this section will appear as Singer et al., „The Effects of Incentives on Response Rates in Face-to-Face and Telephone Surveys“, in a forthcoming issue of the *Journal of Official Statistics*. We would like to thank the anonymous reviewers of that journal, as well as associate editor Nancy Mathiowetz, for their helpful comments.

2) A handful of letters were returned as undeliverable. These did not cluster in any one experimental condition, and are included in the denominator.

3) As noted earlier, some students selected for the sample lived in the same dormitory. As a precaution, we asked, in the course of the telephone survey, whether or not the respondent knew anyone who had been selected for the survey. Eighty respondents--16.2% of the interviewed sample, and 12.8% of the original sample--claimed they did. When the responses of these students are excluded, the attitudinal results do not change

in any important way. We cannot, of course, analyze the effect of this variable on participation in the survey.

4) We stipulated that interviewers were to call back at least 4 times, at different times of the day and on different days, and to make one attempt to convert refusals by those who had not returned the card.

5) In a mail survey, follow-up mailings would serve a similar function.

6) What is considered „equitable“ today may, as a result of exposure to different practices, come to be perceived as „inequitable“ in the future, and vice versa. Thus, the practices of some organizations with respect to incentives may affect the climate in which all surveys are carried out. We return to this topic below.

7) Households who received the \$5 incentive were significantly more likely to respond than those who did not; the difference in response rates was 8 percentage points.

8) The questionnaire and the return address on letters identified the research organization as the Program on Public Opinion and Health Policy at the Survey Research Center, University of Michigan. Hence, respondents may not have identified the mail survey with the organization that had interviewed them more than a year earlier.

9) The number of respondents to the original Detroit Area Study was 451; the follow-up request was sent to 433 of those respondents (18 respondents were omitted because we were unable to locate an address for them). Completed questionnaires were returned by 202 respondents, 4 respondents refused to participate, and 172 respondents did not return questionnaires; 53 questionnaires were undeliverable, and 2 respondents were deceased or could not be contacted for other reasons. The completion rate, as defined earlier, was 53.4%.

10) For a fuller discussion of the research reported here, see Eleanor Singer, John Van Hoewyk, and Patricia Maher, Does the Payment of Incentives Create Expectation Effects? *Public Opinion Quarterly*, Summer 1998. We would like to thank Vincent Price and three anonymous reviewers for their helpful comments on the manuscript.

References

Adams, J. Stacy, 1965: Inequity in Social Exchange. Pp. 267-99 in *Advances in Experimental Social Psychology*, Vol. 2, ed. Leonard Berkowitz. New York: Academic.

Adams, J. Stacy and Sara Freedman, 1976: Equity Theory Revisited. Pp. 43-58 in *Advances in Experimental Social Psychology*, Vol. 9, ed. Leonard Berkowitz and Elaine Walster. New York: Academic.

Cialdini, Robert B., 1988: *Influence: Science and Practice*. Glenview, Il.: Scott, Foresman.

- Church, Allan H., 1993: Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta Analysis. *Public Opinion Quarterly* 57: 62-79.
- Couper, Mick P., Eleanor Singer, and Richard A. Kulka, 1998: Nonresponse in the 1990 Census: Politics, Privacy, and/or Pressures. *American Journal of Politics*.
- Davis, Kingsley, 1948: *Human Society*. New York, Macmillan.
- De Heer, W.F. and Israëls, A.Z., 1992: Nonresponse Trends in Europe. Paper presented at the Joint Statistical Meetings of the American Statistical Association, Boston.
- Festinger, Leon, 1957: *A Theory of Cognitive Dissonance*. Stanford, Stanford University Press.
- Gouldner, Alvin W., 1960: The Norm of Reciprocity. *American Sociological Review*, 25:161-78.
- Groves, Robert M. and Mick P. Couper, 1996: Household-Level Determinants of Survey Nonresponse. In M.T. Braverman and J.K. Slater, eds., *Advances in Survey Research. New Directions for Evaluation*, No. 70. San Francisco: Jossey-Bass.
- Groves, Robert M., Eleanor Singer, Amy Corning, and Ashley Bowers, 1997: The Effects on Survey Participation of Interview Length, Incentives, Differential Incentives, and Refusal Conversion: A Laboratory Approach. Paper presented at the meeting of the American Association for Public Opinion Research, Norfolk, VA.
- Homans, George C., 1961, 1974: *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace, Jovanovitch.
- Hyman, Herbert H. and Eleanor Singer, eds. 1968: *Readings in Reference Group Theory and Research*. New York: Free Press.
- Hyman, Herbert H., 1942: The Psychology of Status. *Archives of Psychology* No. 269.
- Mathiowetz, Nancy A., Theresa J. DeMaio, and Elizabeth A. Martin, 1991: Political Alienation, Voter Registration and the 1990 Census. Paper presented at the annual conference of the American Association for Public Opinion Research, Phoenix, AZ, May.
- Merton, Robert K. and Alice S. Rossi, 1950: Contributions to the Theory of Reference Group Behavior. Pp. 40-105 in *Continuities in Social Research: Studies in the Scope and Method of „The American Soldier,“* ed. Robert K. Merton and Paul F. Lazarsfeld. Glencoe, Ill.: Free Press.
- Newcomb, Theodore M., 1953: An Approach to the Study of Communicative Acts. *Psychological Review* 60: 393-404.
- Shepelak, Norma J. and Duanae F. Alwin, 1986: Beliefs about Inequality and Perceptions of Distributive Justice. *American Sociological Review* 51: 30-46.

-
- Singer, Eleanor, 1981: Reference Groups and Social Evaluations. Pp. 66-93 in *Social Psychology*, ed. Morris Rosenberg and Ralph H. Turner. New York: Basic Books.
- Singer, Eleanor, Nancy Gebler, Trivellore Raghunathan, John Van Hoewyk, and Katherine McGonagle. Forthcoming. The Effects of Incentives on Response Rates in Face-to-Face and Telephone Surveys. *Journal of Official Statistics*.
- Singer, Eleanor, Nancy Gebler, John Van Hoewyk, and Jan Brown, 1997: Does \$10 Equal \$10? The Effect of Framing on the Impact of Incentives. Paper presented at the meeting of the American Association for Public Opinion Research, Norfolk, Va.
- Singer, Eleanor, John Van Hoewyk, and Patricia Maher, 1998: Does the Payment of Incentives Create Expectation Effects? *Public Opinion Quarterly*, forthcoming.
- Smith, Tom W., 1995: Trends in Non-Response Rates. *International Journal of Public Opinion Research*, 7: 157-169.
- Stouffer, Samuel A., et al., 1949: *The American Soldier*. Princeton, NJ: Princeton University Press.
- Suls, J. M. and R. L. Miller, eds., 1977: *Social Comparison Processes*. New York: Wiley.
- Toennies, Ferdinand, 1955: *Community and Association (Gemeinschaft und Gesellschaft)*. London: Routledge and Kegan Paul.
- Walster, Elaine, Ellen Berscheid, and G. William Walster, 1976: New Directions in Equity Research. Pp. 1-42 in *Advances in Experimental Social Psychology*, Vol. 9, ed. Leonard Berkowitz and Elaine Walster. New York: Academic.
- Walster, Elaine, G. William Walster, and Ellen Berscheid, 1978: *Equity: Theory and Research*. Boston: Allyn and Bacon.
- Yu, J. and H. Cooper, 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. *Journal of Marketing Research* 20: 36-44.

DIE MESSUNG VON EINKOMMEN IM NATIONALEN UND INTERNATIONALEN VERGLEICH

*JÜRGEN H.P. HOFFMEYER-ZLOTNIK UND UWE WARNER**

Das Erfassen von Einkommen wird in der Umfrageforschung auf sehr unterschiedliche Art durchgeführt, unterschiedlich hinsichtlich der Definition des Begriffs, der Operationalisierung der Fragestellung, des Umfangs der Abfrage und den Strategien zur Reduktion von Non-Response. Verantwortlich hierfür ist einerseits die Forschungsfrage (welchen Stellenwert hat das Einkommen für die Analyse) und andererseits der „Respekt“, den der Forscher und seine Interviewer vor dieser Frage haben (welches Antwortverhalten wird bei der Abfrage von Einkommen unterstellt). Bewegt man sich auf der Ebene von internationalen Vergleichen, so sind zusätzlich kulturelle Unterschiede in der Handhabung von Statistik und Umfragen zu berücksichtigen.

Ziel dieser Ausführungen kann es nicht sein, das einzig wahre Instrument für die Messung von Einkommen in der Umfrageforschung - am besten international gültig - darzustellen: Es gibt weder ein einziges noch ein wahres Meßinstrument. Ziel ist es, dem Forscher aufzuzeigen, wie die Wirklichkeit in der Einkommensmessung aussieht und was national und international vergleichend bei der Einkommensmessung zu berücksichtigen ist. Hierzu werden wir unterschiedliche Operationalisierungen für die Erfassung von Einkommen vorstellen und die Punkte herausarbeiten, die der Forscher bei einer vergleichenden Analyse von Einkommen berücksichtigen sollte.

The measurement of income differs in many surveys, with respect to the questionnaires, the definitions of income, and the reduction of non-response. This depends on researchers' specific interest in income topics and the answering behavior of the respondents. The paper looks at different strategies for measuring income on the basis of a number of national and international questionnaires. First we present definitions of incomes and their compositions, then consider the operationalization of income. This is followed by national and international differences in income questions. The concept of household is of importance for income research and we consider several household definitions. International income comparisons are made in the last part of our paper.

1. Einleitung

Das Erfragen des Einkommens in der Umfrageforschung hat unterschiedliche Ziele:

- Die Marktforschung ist an der Kaufkraft eines Haushalts interessiert. In den Umfragen kann das Einkommen relativ grob erfaßt werden. Wirklich interessiert der Betrag, der in den Konsum einfließen kann. Dieses ist eine Größe, die so in Umfragen bestenfalls sehr aufwendig erfragt werden kann. Zur Erfassung des möglichen Konsumpotentials benötigt man entweder ein Erfassen von Einnahmen und Ausgaben über einen gewissen Zeitraum oder die Kenntnis über die speziellen Konsummuster spezifischer Zielgruppen (beides ermittelt die Marktforschung mit Hilfe von Paneln). Was in den Umfragen gemessen wird, führt zu einer groben Klassifikation von Konsumgruppen.
- Ein Teil der Sozialforscher ist am Einkommen als einem zentralen Indikator für soziale Ungleichheit interessiert. Denn auch über Einkommen definiert sich die Position einer Person bzw. eines Haushalts in der gesamtgesellschaftlichen Hierarchie von sozio-ökonomischem Status. Unter diesem Analyseaspekt benötigt auch der Sozialforscher in der Regel nicht die Kenntnis eines exakt gemessenen Einkommens. Ihm reicht im Prinzip eine Größenordnung von Individual- oder Haushaltseinkommen. Der sich als Unterschätzung ausweisende Meßfehler kann hinsichtlich Größenordnung und Zusammensetzung beschrieben werden.
- Die ökonomische und sozio-ökonomische Forschung ebenso wie die amtliche Statistik benötigen ein möglichst exakt gemessenes Befragteinkommen. Hier dient das Einkommen zur Analyse der wirtschaftlichen Situation und es steht die Einkommensverteilung und Einkommensentwicklung im Mittelpunkt der Analyse. Dabei geht es allerdings nicht nur um einen Indikator für Wohlstand und dessen Zu- oder Abnahme (oder auch um einen Indikator für Armut), sondern auch um eine Analyse der Einkommensarten, sowie deren einzelner Komponenten in Umfang und Bedeutung.

Die Erhebungseinheit bei der Einkommenserfassung in der Umfrageforschung ist die Befragungsperson, die als Experte über sich und/oder eine Gruppe, in der Regel den Haushalt, Auskunft gibt. Entsprechend dem Analyseziel werden zu einem Haushalt eine Befragungsperson oder alle (zur Grundgesamtheit zählenden) Mitglieder eines Haushalts in den Prozeß der Datenerhebung einbezogen. Die Analyseeinheit sind entweder die befragte Person oder der Haushalt, in den diese eingebunden ist. Nur bei besonderen Fragestellungen wird die Analyseeinheit, mit Implikationen für die Abfrage, geändert, z.B. auf die „Familie“, um hierüber z.B. eine Art von Unterstützungsnetzwerk erfassen zu können.

Die Berechnungseinheit ist entweder das Bruttoeinkommen (vor Abzug von Steuern und Sozialabgaben) oder das Nettoeinkommen (nach Abzug von Steuern und Sozialabgaben). Als dritte Berechnungseinheit wird hin und wieder das „verfügbare Einkommen“ berechnet. Dies ist jener Betrag, ausgehend vom Nettoeinkommen, der nach dem Abzug der Beiträge zur Kranken- und Altersversorgung, der Kosten für Bildung und Ausbildung, der Kosten für Wohnen, regelmäßiger Kosten für Transport, Kommunikation und für Arbeitsmittel einer Analyseeinheit (z.B. einem Haushalt) verbleibt und damit dessen Konsummöglichkeiten beschreibt.

Die Probleme beim Erfassen von Einkommen in der Umfrageforschung sind vielfältig. In der Regel stellt die Erfassung von Einkommen keine Messung, sondern nur eine Schätzung dar. Auch wenn nur der kleinere Teil der Umfrageforscher für ihre spezifischen Fragestellungen exakte Messungen von Einkommen benötigt, so sollten doch auch die Markt-, Media- und Sozialforscher, denen Größenordnungen von Einkommen genügen, Daten erheben, deren Güte akzeptabel ist.

Eine genaue Erfassung des Einkommens erfordert das Führen eines Einnahmen-Ausgaben-Buches, um hierüber, ähnlich einer Steuererklärung, in detail den gewünschten Wert - brutto, netto oder verfügbar - für ein Individuum oder für eine Gruppe von Individuen in ihrer Gruppenzuordnung errechnen zu können. Andere Einkommenserfassungen fragen sehr detailliert die einzelnen Mitglieder eines Haushaltes oder einer anders definierten Gruppe nach deren Einkommensbezug, unterteilt in die unterschiedlichen Kategorien des möglichen Bezugs. Die Masse der Einkommenserfassungen fragt weniger aufwendig ab. In diesem Bereich der Forschung wird Einkommen im Prinzip als eine Gesamtsumme, offen oder kategorial, abgefragt. Die Definition von Einkommen ist in der Regel sehr kurz, manchmal fehlt sie ganz. Die Aufgabe für den Befragten kann von diesem mißverstanden werden. Befragt wird in den „normalen“ Umfragen der Markt-, Media- und Sozialforschung oft nur eine zufällig ausgewählte Person pro Haushalt, die dann als Experte für den Haushalt zum Haushaltseinkommen befragt wird. Ist die befragte Person nicht in der Lage oder willig, die Frage nach dem Haushaltseinkommen zu beantworten, so wird in einer Reihe von Umfragen der Markt- und Mediaforschung (und in deren Gefolge auch der Sozialforschung) der Interviewer angewiesen, nach Augenschein das Haushaltseinkommen des befragten Haushalts zu schätzen. Niemand aber hinterfragt die Expertise des in beiden Fällen das Haushaltseinkommen schätzenden Experten. Im Prinzip kann man dem Befragten nur bei der Frage nach dem eigenen Einkommen Expertenstatus unterstellen, obwohl auch in diesem Fall das Abfragen einer Gesamtsumme nicht die Sicherheit bietet, daß in dieser alle relevanten Teilsommen enthalten sind. Noch weniger kann man von allen Personen eines Haushalts erwarten, daß alle immer

über jede bei irgend einem der anderen Haushaltsmitglieder eingehende Mark informiert sind.

Probleme bei der Erfassung von Haushaltseinkommen über jeweils einen Befragten pro Haushalt ergeben sich über voneinander abweichende Definitionen von „Einkommen“ (was alles zählt dazu?) und von „Haushalt“ (wer alles zählt dazu?). Weitere Probleme entstehen über ein personenspezifisches Unvermögen, die Frage nach dem Haushaltseinkommen zu verstehen (Unverständnis der Aufgabenstellung) oder zu beantworten (Unkenntnis der Summen). Und eine personenspezifische Tendenz zur Verweigerung der Beantwortung der Frage(n) nach dem Haushaltseinkommen macht das Analysieren eines Durchschnittseinkommens noch problematischer. All diese Probleme, die das Abschätzen der Güte der Fragen zum Einkommen in einer Umfrage schon schwierig gestalten, potenzieren sich beim Vergleich über unterschiedliche Umfragen.

Nach diesem Blick auf die nationale Ebene wird die Schwierigkeit eines Einkommensvergleichs auf der internationalen Ebene noch deutlicher, denn hier kumulieren sich Eigenheiten und Probleme der nationalen Abfragen. Wir werden im Folgenden die unterschiedlichen Arten der Einkommensabfragen darlegen und national wie international Umfragedaten miteinander vergleichen, um hierüber die Probleme der Vergleichbarkeit zu diskutieren.

2. Definition und Zusammensetzung von „Einkommen“

2.1 Definition von Einkommen

Unter „Einkommen“ versteht man die einer Person, einer Gesellschaft oder einem Verein aus dessen/deren Teilnahme am Wirtschaftsprozeß in einem bestimmten Zeitraum (Tag, Monat, Jahr) zufließende Mittel. Dieses ist in der Regel ein Geldeinkommen. Naturaleinkommen, sogenannte „Deputate“, findet man im klassischen Sinn noch in der Landwirtschaft, hierzu sind heute allerdings auch kostenlose oder günstige Nutzungsrechte einer Wohnung, eines Autos, von Kommunikationstechnologie oder von Arbeitsgeräten und/oder -räumen zu zählen.

Die Summe der Einkommen aller Personen, die zu einem Haushalt zählen, ist das Haushaltseinkommen. Hinzu kommen haushaltsbezogene Zahlungen staatlicher Transferleistungen. Das Haushaltseinkommen ist neben dem persönlichen Einkommen der zweite zentrale Einkommens-Indikator für die Markt- und Sozialforschung. Im Haushaltseinkommen werden in den meisten Ländern die Nettoeinkommen aller zum jeweiligen privaten Haushalt zählenden Personen addiert. Unter Nettoeinkommen ist das um die Einkommens- und Vermögenssteuern und die Pflichtbeiträge zur Sozialversicherung

(ggf. auch um die Beiträge für private, kommunale und staatliche Zusatzversorgungskassen) verminderte Bruttoeinkommen zu verstehen, das aus den Einkünften aus unselbständiger Tätigkeit, aus Unternehmertätigkeit, aus Vermögen, aus Vermietungen und Verpachtungen und aus öffentlichen und privaten Einkommensübertragungen herrührt. Damit umfaßt das Einkommen sowohl regelmäßige als auch unregelmäßige, temporäre oder saisonabhängige, sowie einmalige Einkünfte aus beruflicher und unternehmerischer Tätigkeit, aus Vermögensbesitz und aus öffentlichen und privaten Transferleistungen. Diese Einkünfte müssen nicht nur monetäre Einkünfte sein - unter gewissen Bedingungen sind auch nicht-monetäre Einkünfte zu berücksichtigen.

In der Bundesrepublik Deutschland ist der zeitliche Bezug für die Einkommensermittlung in der Regel ein Monat: Bei abhängig beschäftigten Erwerbstätigen wird in der Abfrage oft ein konkreter Monat zugrunde gelegt. Bei unregelmäßigem Einkommen und bei Einkommen aus unternehmerischer Tätigkeit wird in der Regel der monatliche Durchschnitt über den Verlauf eines Jahres erfaßt.

2.2 Wie setzt sich Einkommen zusammen

Eine erste grobe Untergliederung von Einkommen unterscheidet in vier Oberbegriffe, hinter denen die ganze Breite des Spektrums von Einkommen steht:

- Einkommen aus Erwerbstätigkeit: Arbeitseinkommen und Einkommen aus Unternehmergeinn
- Einnahmen aus Einkommensübertragungen
- Einkommen aus Besitz und Vermögen
- Einkommen über Deputate und/oder Sachleistungen

Nicht zum Einkommen zählen Einnahmen aus dem Verkauf im Haushalt gebrauchter Waren, Einnahmen aus der Erstattung von Ausgaben für geschäftliche oder dienstliche Zwecke und Rückvergütungen.

Eine Untergliederung der vier Oberbegriffe präzisiert diese am deutschen Beispiel und zeigt die Vielfalt der unterschiedlichen Einkommensarten und möglichen Einkommensquellen auf und gibt damit einen Überblick über die Komplexität einer präzisen Abfrage von Einkommen:

- Einkommen aus Erwerbstätigkeit: Einkommen aus unselbständiger Arbeit (Lohn, Gehalt, Ausbildungsvergütung, Aufwandsentschädigung). Hinzuzurechnen sind Zahlungen für: Urlaubsgeld, Weihnachtsgeld, 13./14./15. Monatsgehalt, Prämien, Überstundenentgelt, Zuschläge, Trinkgelder, vermögenswirksame Leistungen des Arbeitgebers sowie Lohnersatzzahlungen. An dieser Stelle ist auch schon an die

Deputate zu denken. Nicht vergessen werden dürfen Einkommen aus einer zweiten Beschäftigung oder aus nebenberuflicher Tätigkeit.

- Einkommen aus Unternehmergewinn. Einkommen aus Gewerbebetrieb: Der Gewinn eines Gewerbebetriebes wird vereinfacht definiert als Überschuß der Betriebseinnahmen über die Betriebsausgaben. Hinzu kommen Einkünfte aus Gesellschaftergewinnanteilen, sowie Honorare und Aufwandsentschädigungen.
- Für die anderen selbständigen Gruppen der Land- und Forstwirte und die freiberuflich Tätigen sind entsprechend „Einkünfte aus Land- und Forstwirtschaft“ und „Einkünfte aus selbständiger Arbeit“ zu erfassen. Hinzuzurechnen sind Zahlungen über Prämien und Einnahmen aus Einkommensübertragungen.
- Einnahmen aus Einkommensübertragungen: Renten der gesetzlichen Renten- und Unfallversicherung, öffentliche Pensionen, Renten der Kriegsopferversorgung: Altersrenten (Ruhestand, Vorruhestand), Alterspensionen, Renten im Falle von Erwerbsunfähigkeit oder Unfall, Renten für Hinterbliebene, Waisen, Kriegsversehrte und ähnliches mehr.
- Sozialhilfe, Kleidergeld, Sachbeihilfe.
- Arbeitslosengeld, Arbeitslosenhilfe.
- Andere öffentliche Transfers wie: Kindergeld, Familienbeihilfen, Ausbildungsbeihilfen, Wohngeld und anderes mehr.
- Krankengeld, Krankenhaustagegeld, Pflegegeld, Mutterschaftsgeld, Sachbeihilfen, Behindertengeld (nicht in Deutschland).
- Übertragungen von Unternehmen und privaten Organisationen ohne Erwerbscharakter, u.a. Spenden (von caritativen Vereinigungen in Sachen und Geld).
- Übertragungen von privaten Haushalten wie: Alimente, Altenteil; aber auch: Taschengeld und Zuwendungen zum Lebensunterhalt von: Großeltern, Eltern, (Partner), Geschwistern, Kindern.
- Einmalige Einkommenübertragungen, wie z.B. Rückerstattungen von Einkommens- und Vermögenssteuer, aber auch: Schenkungen und Erbschaften.
- Einkommen aus Besitz und Vermögen: Zinsgewinne, Aktiengewinne, Gewinne aus Dividenden, Wertpapieren, aus Fonds, aus Beteiligungen und aus Lotterien, sowie Gewinne aus Hypotheken, Schulden, Renten oder aus Darlehen, Anleihen und Guthaben.
- Einkommen aus Verpachtungen und Vermietungen, einschließlich einer unterstellten Miete aus Eigentümerwohnungen, abzüglich der Aufwendungen für die Unterhaltung von Grundstücken und Gebäuden und für Schönheitsreparaturen der Eigentümerwohnungen.
- Deputate und Sachleistungen: Werkwohnung, Dienstwohnung; Dienstfahrzeug; Jahreswagen; Naturaldeputate; Nutzungsrechte von Werkzeugen, Räumen, Werkstatt, Acker, Kommunikationstechnologie; Erfolgsprämien in Sachleistungen.

Betrachtet man diese Vielfalt der möglichen Einnahmequellen, so zeigt sich sehr schnell, daß eine alles berücksichtigende, genaue Abfrage von Einkommen einer

Einkommenssteuererklärung sehr nahe kommen muß. Denn: Im Prinzip müßte man jeden einzelnen Unterpunkt separat für alle Personen im Haushalt abfragen. Machbarer ist das Erfassen von Summen in größeren Kategorien, auch auf die Gefahr, daß kleine, oft einmalige Summen dabei vergessen werden. Eine ziemlich sichere Methode ist das Erfassen von Einkommen über ein Einnahmen-Ausgaben-Buch, wobei zwei Probleme blieben:

a) Welche Ausgaben stellen Werbungskosten dar und welche nicht? b) Wie sind Deputate und Sachleistungen zu monetarisieren? Solch aufwendige Abfragen, die eher einen speziellen Survey zur Einkommensmessung darstellen als einer Einkommensabfrage in einer allgemeinen Bevölkerungsumfrage nahekommen, sind in der Regel in der Umfrageforschung nicht erhebbar.

3. Zur Operationalisierung der Abfrage von Einkommen

3.1 Ordnung unterschiedlicher Einkommenserfassung

Wie schon oben erwähnt setzen unterschiedliche Analyseziele eine unterschiedlich detaillierte Abfrage der Einkommen/Einkommensverhältnisse von Personen/Haushalten voraus: Eine Ermittlung der Kaufkraft eines Haushaltes im Zusammenhang mit Marktforschung setzt eine andere, weniger aufwendige Operationalisierung voraus als das Ermitteln von Einkommensverteilungen und -entwicklungen für die ökonomische oder sozio-ökonomische Forschung. So reicht dann auch der Spielraum für die Abfrage von Einkommen von einer sehr „freien“ Abfrage von Einkommen mit oft nur einem einfachen Kategoriensystem für Haushaltseinkommen, das einer Referenzperson vorgelegt wird, über eine detailliertere Definition der Aufgabe mit umfangreicherer Abfrageanweisung (mit definierten Einkommensstypen, Zeiträumen des Einkommensbezugs und eventuell festgelegten Bevölkerungsgruppen) bis zu einer für alle Personen eines Haushaltes Teileinkommen addierenden, detaillierten Abfrage.

Die an Einkommensverteilungen und -entwicklungen interessierten Abfragen (hierunter fallen auch die meisten Abfragen der amtlichen Statistik) werden entweder als Querschnitt, wie derzeit der deutsche Mikrozensus und der amerikanische Current Population Survey, March Supplement, oder als Längsschnitt, wie das Sozio-ökonomische Panel (SOEP), die Panel Study of Income Dynamics oder das Europäische Haushaltspanel (European Community Household Panel, ECHP) konzipiert. Die am tiefsten gegliederte Erfassung von Einkommen über eine Zeitperiode in der Regel von einem Jahr stellt das Führen eines Einnahmen-Ausgaben-Buches dar, wie z.B. bei der Einkommens- und Verbrauchsstichprobe der amtlichen Statistik, die eine zwölfmonatige Auflistung von Einnahmen und Ausgaben in einem Tagebuch, davon 1 Monat en detail als Feinanschrei-

bung, den Rest der Zeit z.T. in Kategorien als verkürzte Anschreibung, von den Probanden fordert. Ebenfalls eine tiefe Staffelung bietet das Verfahren des Kalenders, wie er z.B. vom ECHP benutzt wird. Der Kalender deckt meistens eine Jahresperiode ab und erfaßt die Einnahmen (und Ausgaben) monatsweise.

3.2 Elemente der Operationalisierung

Bevor einige zentrale Studien bzw. Studientypen auf deren Operationalisierung der Einkommensabfrage untersucht werden, müssen zunächst die für die Abfrage wichtigen Elemente der Operationalisierung von Einkommen aufgezeigt und diskutiert werden.

Da bei der Einkommenserfassung das Einkommen aus Erwerbstätigkeit für die Mehrheit der Haushalte eine zentrale Rolle spielt, und da bei dessen Abfrage die Art der Beschäftigung, ob abhängig beschäftigt oder unabhängig/selbständig tätig, sich auf die Fragestellung auswirkt, ist vor einer Einkommenserfassung die Art der Erwerbstätigkeit zu klären. In einem weiteren Schritt ist zu klären, ob das Bruttoeinkommen (vor Abzug von Steuern und Sozialversicherungsbeiträgen) oder das Nettoeinkommen erfaßt werden soll. Dieses ist gerade im internationalen Vergleich wichtig zu beachten, da die diesbezüglichen Standards der Abfrage staatenabhängig sehr unterschiedlich sind.

Ein weiteres Element der Abfrage stellt der zeitliche Bezugsrahmen dar: Wird der letzte Monat, der Durchschnittsmonat oder ein anderer Zeitrahmen, wie z.B. die Woche oder das Jahr zugrunde gelegt? Lohn und Gehalt werden in Deutschland im Regelfall monatlich überwiesen - aber der letzte Monat weicht vom Durchschnittsmonat ab, da es zusätzlich eine ganze Reihe von Einnahmen in einem Haushalt gibt, die nicht monatlich, sondern oft nur einmalig pro Jahr dem Bezieher zufließen. Gewerbebetriebe, Landwirte und Freiberufler können mit einer Monatsbasis der Abfrage wenig anfangen - für diese Gruppen ist die Jahresbasis, eventuell transformiert zum Durchschnittsmonat sinnvoll. Bei spezifischer Fragestellung sollte man allerdings nicht vergessen, daß es für kleine Gruppen auch noch Tagelohn oder Wochenlohn gibt und daß bestimmte Transferleistungen im Wochenrhythmus zugeteilt werden. Betrachtet man vergleichend die internationale Ebene, so sind die Abweichungen vom „letzten Monat“ häufig gegeben. Hier ist es wichtig, sich im Vorfeld mit den Gepflogenheiten in den Vergleichsländern vertraut zu machen, besser noch: sich mit den Forschungspartnern in den anderen Ländern über dieses Problem abzustimmen.

Schließlich ist es wichtig, der Befragungsperson den vom Forscher gewünschten Begriff von Einkommen zu definieren und die Aufgabe so exakt zu stellen, daß diese bewältigbar ist. D.h., je detaillierter Einkommen erhoben werden soll, desto ausführlicher muß angegeben werden, welche Einkommensarten zu berücksichtigen und zu addieren sind. Einkommen aus Erwerbstätigkeit plus Transfereinkommen (welche

dürfen auf keinen Fall vergessen werden?) plus Einkommen aus Besitz und Vermögen; und sollen auch nicht-monetäre Einkommen berücksichtigt werden? Was soll subtrahiert werden? Die Definition des Nettobetrags erfordert in der Bundesrepublik den Abzug von „Steuern und Sozialversicherungsbeiträgen“ - wie sind diese definiert?

Entscheidend für die Beantwortbarkeit sind auch die Technik und der Umfang der Abfrage, deren Detailliertheit in Vorgaben und Erläuterungen, die Wahl einer offenen oder einer kategorialen Abfrage zur Erfassung errechneter Beträge. Schließlich ist noch die Entscheidung zu treffen, ob die Einkommensangabe im Fragebogen ein „objektiv gemessener“ Wert sein soll: Der Befragte gibt als Experte über seine Einkünfte, tief gestaffelt oder weniger tief gestaffelt Auskunft mittels einer Einnahmen-Ausgaben-Buchführung (EVS) oder über die Steuerunterlagen (wie z.B. in Frankreich üblich) oder anhand ad hoc hinzugezogener Unterlagen oder aus dem Gedächtnis.

Alternativ zum „objektiv gemessenen“ Wert steht der „subjektiv erfaßte“ Wert, der nur dann als zu erhebender Wert gewählt werden darf, wenn die Forschungsfrage nicht an der Einkommensverteilung und -entwicklung interessiert ist. Hierbei werden Teilbeträge und deren Addition zu einer Gesamtsumme durch die Befragungspersonen ohne Hilfestellung aus Unterlagen erinnert und errechnet. Oft wird allerdings nur ein „subjektiv geschätztes“ Einkommen erfaßt, vor allem dann, wenn eine Referenzperson zum ganzen Haushalt befragt wird. Dieser Wert gibt eine relativ unverbindliche Größenordnung wieder und ist damit nur bei entsprechenden Fragestellungen sinnvoll. Als Alternative zum geschätzten Wert durch den Befragten findet man bei Umfragen der Markt- und Mediaforschung im Falle der Verweigerung der Angabe durch den Befragten auch einen durch den Interviewer geschätzten Wert. Diese Schätzung basiert im Wesentlichen auf Informationen der Abfrage von Beruf oder Lebensstilvariablen und auf dem persönlichen Eindruck des Interviewers, den dieser vom Befragten und dessen räumlicher und sozialer Umgebung hat. Ob ein solcher Wert in der Sozialforschung akzeptiert werden kann, muß jeder Forscher entsprechend seiner Fragestellung entscheiden. Grobe Größenordnungen sind auch über Interviewerschätzungen zu erhalten. Im internationalen Vergleich reicht der Marktforschung oft auch eine Auflistung über den Besitz langfristiger Konsumgüter. Auch diese Bestimmung der Kaufkraft ist für die Sozialforschung nicht immer akzeptabel und für die Messung sozialer Ungleichheit unbrauchbar, da spezifische Konsumgüter einerseits kulturabhängig sind (unter einer Kaffeemaschine versteht man in Italien etwas anderes als in Deutschland) und andererseits sich sehr in der Qualität unterscheiden (Fotoapparate, Videorecorder oder Computer gibt es in sehr unterschiedlichen Qualitäts- und damit Preisabstufungen).

Das letzte Element der Operationalisierung ist die Bezugseinheit der Datenerfassung: Ist das Individualeinkommen die Erhebungseinheit oder soll das Einkommen einer definierten Gruppe, z.B. eines Haushalts oder einer Familie, erfaßt werden? Was ist die Analyseebene: Individuum oder Gruppe bzw. Haushalt?

3.3 Die Operationalisierung von Einkommen in konkreten Studien

Unter diesem Punkt werden wir die Einkommensfragen aus einer Reihe von Studien aus der amtlichen Statistik, aus nationalen Umfragen und aus internationalen Umfragen vorstellen. Eine Auflistung einzelner Surveys nach den genannten Kriterien befindet sich im Anhang.

Abfragen von Einkommen in der amtlichen Statistik, national und international

Der *Mikrozensus* erfragt für abhängig Beschäftigte das Einkommen abzüglich Lohnsteuer, Kirchensteuer, Sozialversicherungs- u.ä. Beiträge. Selbständige werden über das Unternehmereinkommen erfaßt. Der zeitliche Bezug der Antworten zu den Einkommensfragen ist ein festgelegter Stichmonat. Die Einkommensstypen sind ausführlich aufgelistet: Sie umfassen Lohn, Gehalt, Unternehmereinkommen, diverse Renten, Pensionen, Einkünfte aus Vermietung, Verpachtung, Gratifikationen, vermögenswirksames Sparen, Vorschüsse, Werkwohnungsmiete, differenziert gelistete öffentliche und private Transferleistungen. Das Nettoeinkommen wird kategorial in 18 Kategorien abgefragt, hinzu kommen die Gruppen der mithelfenden Familienangehörigen, der Landwirte und jene Personengruppe, die über kein Einkommen verfügt. Die Einnahmen werden als monetäre Beträge abgefragt. Schätzungen von Angaben sind nicht vorgesehen. Die Erläuterungen weisen den Befragten darauf hin, was er dem Nettoeinkommen zuzurechnen hat. Neben der Frage nach dem Nettoeinkommen wird nach der Quelle des überwiegenden Lebensunterhalts (Erwerbstätigkeit, Rente, etc.) gefragt.

Noch tiefer gestaffelt erhebt die *Einkommens- und Verbrauchsstichprobe* das Einkommen. Für Selbständige werden die Angaben der letzten Steuererklärung erfaßt und für Nicht-Selbständige werden Beträge abzüglich der Steuern und Pflichtbeiträge zur Sozialversicherung notiert. Protokolliert werden neben einem Grund- und einem Schlußinterview in einem Haushaltsbuch über einen Monat alle Einnahmen und Ausgaben en detail als „Feinanschreibung“ und über zwölf Monate in einer verkürzten Anschreibung, Einnahmen und Ausgaben kategorisierend. Die Einkommensstypen sind ausführlich aufgeführt und werden im Haushaltsbuch unter 55 Items notiert.

Etwas weniger tief gestaffelt stellt das *European Community Household Panel* (ECHP) die Fragen sowohl zum Netto- als auch zum Bruttoeinkommen der Haushaltsmitglieder.

Daneben wird eine Referenzperson zum Haushaltsnettoeinkommen befragt. Die Brutto-bezüge sind als „vor Steuern und Sozialabgaben“ beschrieben, die Nettoeinkommen als „the take home pay“. Selbständig tätige Personen werden nach dem „profit“ nach Abzug aller Ausgaben und Löhne als Brutto befragt. Die Einkommen werden retrospektiv für jeden Monat eines Referenzjahres erfaßt und zusätzlich für den Monat, in dem das Interview stattfindet. Im ersten Schritt gibt es offene Antwortmöglichkeiten, bei Verweigerung werden die Einkommensfragen mit kategorialen Antwortvorgaben wiederholt. Bei Nichtwissen der genauen Beträge besteht für den Befragten die Möglichkeit, einzelne Angaben zu schätzen. Die Definitionen der Einkommensstypen werden in den Frageformulierungen vorgegeben und in einem Interviewerheft erläutert.

Abfragen von Einkommen in der deutschen Umfrageforschung

Als Beispiel für eine Umfrage, deren Analyseziele Einkommensverteilungen und -entwicklungen darstellen, steht das *Sozio-ökonomische Panel (SOEP)*. Das Sozio-ökonomische Panel definiert Selbständige über den „Gewerbebetrieb“ oder über „selbständige Tätigkeit“ und fragt diesen Personenkreis nach Einkünften aus dem Betrieb oder der Tätigkeit laut Einkommenssteuerbescheid. Nicht-selbständige Erwerbspersonen werden gebeten, ihr Nettoeinkommen (Definition: Steuerpflichtiges Einkommen minus Steuern und Sozialabgaben plus Wohngeld, Kindergeld, BaföG und Unterhaltszahlungen) anzugeben. Transferleistungen werden an Hand von Auflistungen als Bruttobeträge abgefragt und in offenen Antworten protokolliert. Die verschiedenen Einkommensstypen sind in den Frageformulierungen genannt und in den Intervieweranweisungen weiter beschrieben. Befragt werden alle Einkommensbezieher eines Haushaltes. Der Prozeß der Einkommensentwicklung ist über das Panel gewährleistet.

Die *Demographischen Standards* für Marktforschung und Sozialforschung, die sowohl zur Analyse von Kaufkraft als auch zur Analyse von sozialer Ungleichheit eingesetzt werden können, erfragen Einkommen aus selbständiger Tätigkeit abzüglich Betriebsausgaben, sowie für abhängig Beschäftigte und Ruheständler: Einkünfte abzüglich Steuern und Sozialversicherungsbeiträgen. Zusätzlich zu berücksichtigen sind Einkünfte aus „öffentlichen Beihilfen, Einkommen aus Vermietung, Verpachtung, Wohngeld, Kindergeld und sonstige Einkünfte“. Die befragte Person, eine pro Haushalt, wird sowohl zu dem eigenen Monatseinkommen als auch zu dem monatlichen Haushaltseinkommen befragt. Die Antwort wird zunächst offen erhoben und bei Verweigerung wird kategorial nachgefragt. Die zum Gesamteinkommen zu addierenden Einkommensstypen sind in der Frageformulierung und in den Erläuterungen benannt.

In die gleiche Kategorie von Umfragen fällt auch der *ALLBUS*. Das Vorgehen der Datenerhebung ist ähnlich. Auch hier wird das Einkommen zunächst offen und bei Verwei-

gerung nachfolgend kategorial erfaßt. Im Unterschied zu den Demographischen Standards werden die verschiedenen Einkommenstypen nicht explizit benannt.

Die *Media-Analyse*, interessiert an einem exakten Erfassen von Kaufkraft, erfaßt wie die beiden vorangegangenen Surveys Monatseinkommen, das „selbst verdient“ ist oder „aus Renten, Pensionen, Mieten, Pacht, Wertpapieren und dergleichen“ stammt; Transferleistungen sind nicht in der Abfrage enthalten. Neben der Befragtenantwort als erstem Schritt und dem Schätzen des Befragten bei Unkenntnis des wahren Wertes als zweitem Schritt, ist bei dieser Umfrage bei endgültiger Verweigerung durch den Befragten als dritter Schritt die Schätzung des Einkommens durch den Interviewer zulässig.

Für die einfachste Art der Einkommenserfassung steht, stellvertretend für viele Abfragen der Marktforschung, die *GFM-GETAS-Großumfrage*, die weder nach Gruppen unterscheidet noch Einkommen differenziert, sondern lediglich ein Kategorienschema vorlegt und den Befragten bittet, hierin den Haushalt zu verorten. Erhoben wird das monatliche Nettoeinkommen in einer Zuordnung, die Rückschlüsse auf eine Größenordnung für Kaufkraft ermöglicht.

Abfragen von Einkommen in der internationalen Umfrageforschung

Auch außerhalb Deutschlands gibt es die drei unterschiedlichen, an den Fragestellungen orientierten Techniken der Operationalisierung der Einkommenserfassung:

Die amerikanische *Panel Study of Income Dynamics* erfaßt die verschiedenen Einkommen differenziert durch die Abfrage jedes möglichen Types pro Person. Bruttoeinkommen aus abhängiger Beschäftigung ist als „before taxes or other things“ gekennzeichnet. Selbständigeneinkommen ist als „share the income from a business“ identifiziert. 48 Fragen pro Individuum plus zusätzliche Unterfragen sind offen formuliert.

Deutlich tiefer gestaffelte Analysen sind in jenen Ländern wie Norwegen, Schweden oder Frankreich möglich, in denen die Einkommensabfrage, auch für die universitäre Sozialforschung, ersetzt werden kann durch einen Datenabgleich mit dem Finanzregister. Der Datenabgleich mit dem Finanzregister geschieht anonymisiert durch die Bearbeitung der Identifikationsnummern der Zielpersonen mit Hilfe eines Zufallszahlengenerators. Dieser Rückgriff auf das Finanzregister steht auch dem Umfrageforscher bei entsprechender Fragestellung zur Verfügung. Daher gibt es außerhalb der Marktforschung in diesen Ländern keine etablierten Instrumente der Sozialforschung zur Einkommenserfassung.

Der us-amerikanische *General Social Survey GSS*, interessiert an der Messung sozialer Ungleichheit, erhebt Bruttoeinkommen „before taxes“ auf jährlicher Basis („last year“). Erwerbseinkommen wird als „earned any money from any job or employment“ und „all

sources“ erfragt. Es gibt kategoriale Antwortmöglichkeiten, aber keine weiter vertiefenden Erklärungen zu den Einkommensfragen.

Marktforschungsumfragen, deren primäres Interesse die Ermittlung der Kaufkraft ist, sehen international entsprechend der deutschen Abfrage aus. Es wird ein Kategorienschema vorgelegt und die Aufgabe lautet, den Haushalt hierin zu verorten. Auf einer entsprechenden Ebene, aber mit anderer Logik, fragt die *ESOMAR*, der Dachverband der europäischen Marktforscher, Kaufkraft nicht monetär, sondern über den Besitz langfristiger Konsumgüter ab: Aufgelistet werden insgesamt zehn Güter, die nach vorhanden/nicht vorhanden erfaßt werden (u.a. Farbfernseher, Zweitwohnung).

Die vergleichende Sozialforschung, hierfür soll das *International Social Survey Programme (ISSP)* angeführt werden, erhebt die Einkommensinformationen von Land zu Land unterschiedlich, gemäß der landesüblichen Abfrage in sozialen Surveys. Die meisten Teilnehmerländer des ISSP stellen die Einkommensfrage zunächst offen und bei Verweigerung mit kategorialer Antwortvorgabe. Ein Vergleich der Länder USA, Großbritannien, Nord-Irland, Republik Irland, Niederlande, Deutschland, Österreich, Italien und Ungarn zeigt: Der zeitliche Bezug der Einkommensangaben ist für die USA und Irland jährlich, sonst monatlich. Deutschland, Österreich und Italien erheben Nettobeträge, die anderen Länder Bruttoeinkommen. Die Einnahmen werden in der jeweiligen Landeswährungen erfragt. Die abgefragten Einkommensstypen, deren Definitionen und die Erläuterungen zu den Fragen sind zwischen den Teilnehmerländern heterogen, zumeist jedoch knapp gehalten. Da auch noch die Definition von Haushalt variiert, ist ein Vergleich schwierig.

3.4 Unterschiede bei der Operationalisierung

Die Auflistung der unterschiedlichen Strategien, Einkommen in Untersuchungen abzufragen, ist bei weitem nicht vollständig und erschöpfend. Aber sie zeigt die Variationsbreite der Erhebungsinstrumente auf. Diese unterschiedlichen Abfragemethoden haben Auswirkungen auf die übermittelten Informationen und auf die Analyseergebnisse. In der vergleichenden Sozialforschung haben sich für die Erhebung von Einkommensinformationen folgende Operationalisierungsvorschläge durchgesetzt:

Die im *ESSPROS Manual* (Eurostat 1995) aufgelisteten und beschriebenen Typen der Einkommen legen ein Hauptaugenmerk auf die Einkünfte aus den sozialen Sicherungssystemen. Auf dieser Systematik basiert die Einkommensabfrage des ECHP. Eine weitergehende Standardisierung, die dem internationalen Vergleich von Einkommen dienen soll, findet sich bei der *Luxembourg Income Study (LIS)* (vgl. Warner 1994). Die LIS-Datenbank verfügt über zwölf Einkommensmerkmale für Individuen und über 51

Einkommensaggregate für Haushalte. Einen weiteren Zugang zu detaillierten monetären Informationen, sowohl auf der Ebene der Haushalte, wie auch auf der Ebene von Individuen bietet das *Panel Comparative (PACO)*-Projekt mit 66 verschiedenen Einkommensvariablen (vgl. Schmaus o.A.J.). Eine weit verbreitete Vorgehensweise stellt jedoch die länderspezifische Abfrage von Einkommen (entsprechend dem ISSP) dar.

Auf der international vergleichenden Ebene kommen zu den durch eine unterschiedliche Operationalisierung hervorgerufenen Unterschieden der Einkommenserfassung auch „kulturelle“ Unterschiede hinzu, bedingt durch die Unterschiede in den nationalen Wirtschaftssystemen, bei den nationalen Besteuerungssystemen (hohe direkte oder indirekte Steuern, unterschiedliche Möglichkeiten von Steuerabschreibungen) und den nationalen Sozialversicherungssystemen: Wird das Einkommen z.B. brutto oder netto erhoben? Und was bedeutet „netto“? In der Bundesrepublik Deutschland wird der Netto-Betrag über den Abzug von „Steuern und Sozialversicherungsbeiträgen“ definiert. In den meisten anderen Ländern, in denen ein Netto-Betrag erhoben wird, ist dieser definiert allein über den Abzug von „Steuern“. Aber auch Unterschiede durch differierende zeitliche Bezüge sind zu berücksichtigen. Hier gibt es schon das Problem innerhalb einer Kultur im Vergleich zwischen abhängig Beschäftigten und selbständigen Erwerbstätigen: In der Bundesrepublik Deutschland rechnen Arbeiter, Angestellte und Beamte ihr Einkommen auf der Monatsbasis, Selbständige aber auf einer jährlichen Basis, wobei unterschiedliche Berufsgruppen unterschiedlich definierte Wirtschaftsjahre zugrunde legen. Problematischer wird es im internationalen Vergleich, da hier nationale Vorlieben der Messung zu berücksichtigen sind.

4. Nationale und internationale Unterschiede

4.1 Unterschiede im Item-Non-Response

Zum Herabsetzen der bei der Einkommensabfrage hohen Item-Non-Response-Rate werden unterschiedliche Strategien angewandt: Entweder der Forscher versucht durch eine thematische Einbindung der Einkommensabfragen eine spezifische Vertrauensbasis zwischen Interviewer und Befragtem herzustellen, die auch einen nochmaligen Hinweis auf die Vertraulichkeit der Antworten beinhalten kann. Oder der Forscher versucht über eine zweistufige Abfrage, eingangs offen fragend und für die Nichtantworter nachfolgend Kategorien vorlegend, den Non-Response zu reduzieren. Wie gut diese zweistufige Abfrage den Item-Non-Response-Anteil reduziert, zeigt in Tabelle 1 der Vergleich zwischen den beiden Listungen für das ISSP: Der Anteil von gültigen Antworten durch eine zusätzliche kategoriale Abfrage steigt von 55,8 Prozent auf 75,5 Prozent. Eine weitere Möglichkeit der Reduktion von Item-Non-Response ist das

explizite Zulassen von Einkommenschätzungen - bis zur Einkommenschätzung nach Augenschein durch den Interviewer allein. Solch einer Schätzung muß die Verweigerung der Angabe durch den Befragten vorausgehen.

Tabelle 2 zeigt deutlich, daß durch unterschiedliche Techniken der Datenerhebung der Non-Response sich verringern läßt. Allerdings zeigt ein Blick in die letzte Spalte, daß das einfachste Instrument, bestehend aus einer Liste mit Einkommenskategorien und der Überschrift: „Monatliches Haushalts-Netto-Einkommen“ (siehe Anhang, Instrument Omnibusumfrage), mit Blick auf die Antwortbereitschaft der Befragungsperson den geringsten Non-Response erzeugt und daher ein Nachhelfen durch Interviewereinschätzungen nicht nötig hat.

Tabelle 1: Monatliches Haushaltseinkommen: Vergleich zwischen ISSP und SOEP, Erhebungen Deutschland-West 1990 (in Prozent)

	ISSP (offene Abfrage)	ISSP (offen + Kategorie)	SOEP (Individuen addiert)
Gültige Antworten	55,8	75,5	62,8
Verweigert	42,3	19,7	37,0
Unbekannt	1,9	5,2	0,2
10% Dezil	1200	1125	329
20% Dezil	1600	1625	718
30% Dezil	2000	2125	1196
40% Dezil	2500	2625	1775
50% Dezil	2800	2875	2310
60% Dezil	3100	3250	2803
70% Dezil	3600	3750	3344
80% Dezil	4100	4250	4117
90% Dezil	5100	5250	5613

Eigene Berechnung; die Individualeinkommen des SOEP sind für die Wohnbevölkerung auf den Haushalt aggregiert.

Tabelle 2: Stufenweises Abfragen zur Reduktion von Non-Response bei Demographischen Standards (Sozialwissenschaften-Bus), Media-Analyse und GFM-GETAS-Großumfrage, 1993 für Westdeutschland (in Prozent)

Stufen	Sozialwissenschaften-Bus	Media-Analyse	GFM-GETAS Großumfrage
1.	offen 51,7	spontan 46,3	kategorial 92,6
2.	kategorial 25,0	Nachfrage 18,2	

3.		Interviewer- schätzung	35,5	
insgesamt	76,7		100	92,6

Berechnung: von Harder, GFM-GETAS Sozialforschung.

Steigt aber mit dem Aufwand der Abfrage auch die Güte der Ergebnisse? In Tabelle 3 sind die Angaben nach der letzten Stufe der Erhebung der Einkommensdaten für die Haushaltseinkommen in Kategorien dargestellt. Bei der Media-Analyse wurde allerdings unterschieden zwischen der Datenerhebung durch Befragung und dem Gesamtergebnis von Befragtenantwort und Interviewereinschätzung.

Tabelle 3: Haushaltseinkommen: Vergleich zwischen verschiedenen Umfragen (Westdeutschland, in Prozent)

Kategorien in DM	Sozialwissen- schaften-Bus	Media-Analyse		GFM-GETAS Großumfrage	Mikrozensus
		nur Befragten- angaben	zusätzlich Interviewer- schätzung		
unter 1000	1,4	1,6	1,4	1,3	2,8
1000-1499	5,4	5,0	4,2	4,2	
1500-1999	7,1	8,1	7,2	8,1	
2000-2499	11,5	10,9	10,3	12,5	
1000-2499	25,4	25,6	23,1	26,1	21,4
2500-2999	12,1	12,6	12,1	15,6	11,0
3000-3499	11,9	10,9	10,3	12,5	11,9
3500-3999	9,5	11,1	11,4	10,7	9,9
4000-4499	10,0	9,2	9,7	10,1	9,0
4500-4999	7,7	9,2	9,6	7,9	7,5
5000-5999	10,4	8,9	9,7	9,0	10,3
6000 u. mehr	14,0	12,5	13,3	8,3	16,2
insgesamt	100	100	100	100	100

Sozialwissenschaften-Bus 3/93 (ZUMA-Standarddemographie/Demographische Standards), der Media-Analyse 1993 und der GFM-GETAS-Großumfrage 1993 sowie dem Mikrozensus 1993.

Berechnung: von Harder, GFM-GETAS Sozialforschung und Statistisches Bundesamt 1995: 57.

Betrachtet man die Verteilungen in Tabelle 3, so wird deutlich, daß das einfachste Instrument in den Kategorien um das durchschnittliche Haushaltseinkommen am stärksten besetzt ist. Die Enden unten und oben werden im Vergleich zum Mikrozensus deutlich unterschätzt. Die Abfragen der Sozial- und Mediaforschung „Sozialwissenschaften-Bus“ und „Media-Analyse“ bieten, soweit nicht der Interviewer bei der Media-Analyse schätzend eingreift, im Vergleich zum Mikrozensus ein leichtes

Überschätzen der unteren Einkommensgruppen. Vermutlich schlägt sich hier das Problem der einen Referenzperson für einen Haushalt nieder, denn die per Zufallszahlenreihe ermittelten Referenzpersonen haben oft keinen Überblick über den Haushalt - vermutet wird, daß diese, an einem allgemeinen oder ihrem eigenen Einkommen orientiert, zu einem nicht unerheblichen Teil den wahren Wert stark unterschätzen.

Zieht man den unterschiedlichen Unit-Non-Response auf der internationalen Ebene in die Betrachtung der Einkommensabfrage mit ein, so ergibt sich, wie Tabelle 4 zeigt, ein die Nationen differenzierendes Bild. Allein die allgemeinen Antwortraten (Tabelle 4, Spalten 1 und 2) der Haushalte und der Personen differieren von Land zu Land. Die Länder mit Auskunftspflicht (Belgien und Griechenland) haben die höchsten Antwortraten. Das allgemeine Bild wird allerdings eher durch mittlere bis niedrige Antwortraten gekennzeichnet. Hierbei muß beachtet werden, daß in unterschiedlichen Nationen eine unterschiedlich hohe Antwortbereitschaft angetroffen wird. Hinzu kommt allerdings, wie das belgische Beispiel zeigt, eine trotz Auskunftspflicht zwischen unterschiedlichen Erhebungsinstituten differierende Ausschöpfungssrate beim Interviewen. Eine unterschiedliche Stichprobenausschöpfung kann allerdings einen Bias in den Antworten bewirken, denn wir gehen nicht davon aus, daß sich die Ausfälle zufällig verteilen: man kennt einen Bias nach Schicht (Mittelschichtbias), einen nach Alter, nach Lebenszyklus und nach Haushaltsgröße.

Tabelle 4: Antwortraten des europäischen Haushaltspanels, in Prozent

Land	Antwortraten der ...		Antwortausfälle auf Fragen zu ...		
	Haushalte	Personen	Net Income	Income from Capital	Rental Income
Belgien (PSBH)	86	81			
Belgien (CSB)	75	66			
Belgien			3,3	9,5	3,6
Deutschland	48	38	2,7	16,4	0,5
Spanien	67	65	1,2	1,0	2,2
Frankreich	79	75	2,7	66,9	9,3
Griechenland	90	86	0,0	0,0	0,0
Italien	89	88	4,0	2,0	44,7
Luxemburg	40	39	4,0	6,2	3,1
Niederlande	36		1,1	0,0	2,4
Großbritannien	70	67	4,2	5,6	4,4

Antwortpflicht in Griechenland und Belgien; die belgischen Daten wurden von zwei unterschiedlichen Instituten erhoben; die Antwortausfälle (Spalten 4-6) stellen den Durchschnitt dar; für die Niederlande sind die Daten des Panel nicht berücksichtigt. Quelle. Europäische Kommission 1996: 36 und 38.

Ein Blick auf die Spalten 4-6 in Tabelle 4 zeigt, daß zu der allgemeinen Antwortbereitschaft noch ein fragenspezifischer Item-Non-Response hinzukommt, der nationenspezifisch ist und zu weiteren Verzerrungen führen kann. Interessant ist allerdings, daß der Item-Non-Response vor allem in jenen Ländern auftritt, in denen der Unit-Non-Response noch auf niedrigem Niveau liegt. Beide Arten von Antwortausfällen können das Bild der nationalen Einkommenverteilungen verzerren.

Allerdings gibt es auch Länder, in denen Fragen zum Einkommen als sehr heikle Fragen gelten, so daß nur in Ausnahmefällen eine Ermittlung von monetärem Einkommen in allgemeinen Bevölkerungsumfragen der Sozial- und Marktforschung stattfindet. Für diesen Typus von Ländern eignet sich am ehesten die ESOMAR-Abfrage nach dem Besitz langfristiger Konsumgüter zur Verortung von Kaufkraft. Die ESOMAR-Abfrage stellt jedoch im engeren Sinne keine Einkommensabfrage dar und eignet sich schon nicht mehr zur Messung sozialer Ungleichheit.

4.2 Unterschiede in den Daten und Möglichkeiten des Vergleichs

Die durch eine unterschiedliche Fragestellung bedingten Unterschiede in der Operationalisierung der Einkommenserfassung führen zu unterschiedlichen Ergebnissen. In Tabelle 5 werden zwei unterschiedlich aufwendig erhobene Einkommensverteilungen miteinander verglichen. Das deutsche ISSP erhebt Einkommen mit einer zwar Einkommen definierenden, aber globalen Abfrage über eine Referenzperson in zwei Stufen: erst offen und nachfolgend kategorial. Das Sozio-ökonomische Panel (SOEP) erhebt Einkommen über eine Vielzahl von Einzelfragen zu Teileinkommen, gesondert für jede Person im Haushalt. Die so pro Person und pro Einkommensart in feingliederten Instrumenten erhobenen Einzeldaten werden schließlich vom Forscher zum Haushaltseinkommen addiert. Tabelle 5 zeigt, daß die detaillierte Einkommensabfrage des SOEP zu einer stärkeren Verteilung der Nennungen in den oberen und unteren Bereichen eines feingliedrigen Kategorienschemas führt. Je gröber die Abfrage nach dem Haushaltseinkommen und je mehr die Konzentration der Abfrage auf das Arbeitseinkommen abzuheben scheint, desto stärker konzentriert sich die Einkommensverteilung der Haushalte auf die mittleren Einkommenskategorien.

Tabelle 5: Haushaltseinkommen: Vergleich zwischen ISSP und SOEP, Verteilung Deutschland-West 1990 in Kategorien (in Prozent)

Kategorien in DM		ISSP	SOEP
unter	400	0,2	4,5
400	bis 599	0,5	1,4
600	bis 799	1,3	1,2

800	bis	999	2,5	2,0
1000	bis	1249	3,3	1,8
1250	bis	1499	4,0	1,3
1500	bis	3999	69,9	64,1
4000	bis	4499	6,4	4,3
4500	bis	4999	3,9	3,7
5000	bis	5499	2,8	3,9
5500	bis	5999	1,6	3,1
6000	bis	7999	3,6	8,7

Vergleicht man Einkommensverteilungen über Länder, so muß zunächst sichergestellt werden, daß die Abfrage Gleiches mißt. In Tabelle 6 wird das demonstriert, was normalerweise geschieht. Verglichen werden die Daten des ISSP für Deutschland-West, Italien, Großbritannien und die USA. Damit eine Vergleichbarkeit der Beträge möglich wird, werden alle Beträge in ECU umgerechnet. Es stellt sich jedoch die Frage, ob hierüber wirklich eine Vergleichbarkeit gegeben ist, denn hinter den Einkommenserfassungen stehen unterschiedliche Definitionen von Einkommen, vom Erhebungszeitraum und von Haushalt. Zu beachten ist, daß die Abfragen in Deutschland und Italien von weitgehend identischen Definitionen ausgehen: Nettoeinkommen, auf monatlicher Berechnungsbasis, Haushalt wird als Wirtschaftseinheit definiert. Die Abfragen in den USA und in Großbritannien erheben Brutto-Einkommen auf Jahresbasis und definieren den Haushalt über die Familie (USA) bzw. über gemeinsame Mahlzeiten (UK).

Tabelle 6: Haushaltseinkommen in Deutschland-West („netto“), Italien („netto“), Großbritannien („brutto“) und den USA („brutto“), umgerechnet auf Monatseinkommen in ECU (in Prozent)

	Deutschland-W.	Italien	Großbritannien	USA
unter 500 ECU	4,5	5,6	9,9	11,7
500-749 ECU	7,4	5,9	12,3	8,4
750-999 ECU	7,7	10,5	7,9	4,0
1000-1249 ECU	10,3	20,5	6,6	9,9
1250-1499 ECU	8,9	7,4	8,6	4,9
1500-1749 ECU	9,5	6,2	0,0	11,2
1750-1999 ECU	6,5	11,1	8,9	0,0
2000-2249 ECU	6,4	4,7	5,7	8,0
2250-2499 ECU	3,9	3,0	6,6	6,5
2500-2749 ECU	2,8	6,7	0,0	0,0
2750-2999 ECU	1,6	1,5	4,8	8,0

3000-3499 ECU	3,6	5,7	6,3	0,0
3500-3999 ECU	0,0	0,0	6,9	18,5
4000 und mehr	2,0	0,0	0,0	0,0
missing data	25,0	11,2	15,5	9,0
Landeswahrung:				
Mean	3266.5	2455.9	14867.0	29686.8
Median	2875.0	2250.0	13000.0	26250.0
std.Deviation	2049.2	1204.2	9096.7	18445.1

Quelle: ISSP, eigene Berechnung; Italiens Landeswahrung in 1000 Lire.

Tabelle 6 suggeriert zwar einen Vergleich zwischen den vier gelisteten Landern - dennoch ist nicht eindeutig zu identifizieren, mit jeweils welchem Anteil hier a) ein kultureller bzw. nationenspezifischer, b) ein sozio-onomischer und c) ein methodischer Unterschied gemessen wird. Von der Einkommenserfassung am ahnlichsten sind sich Deutschland-West und Italien - auch wenn die Verteilungen voneinander abweichen. Anders sieht es mit den Verteilungen fur Grobritannien und die USA aus. Beide weisen jeweils eine von der ersten Gruppe (Deutschland und Italien) abweichende Struktur auf: Brutto, Jahresbasis, andere Haushaltsdefinition. Bei der Haushaltsdefinition unterscheiden sich letztere beiden selbst noch einmal voneinander. Neben sozio-onomischen Effekten, die es in den Umfragen in der Regel herauszufiltern gilt, spielen hier auch kulturelle oder nationenspezifische Effekte mit, die die Operationalisierung der Abfrage beeinflussen und sich in der sichtbar werdenden Struktur von Einkommen niederschlagen. Operationalisierungsaspekte sind bei der Interpretation solcher vergleichender Tabellen von Bedeutung.

Die Moglichkeiten der Umfrageforscher, Einkommen fur die Analysen im internationalen Vergleich als „reale“ Werte nutzen zu konnen, basieren, neben einer Erfassung nach vergleichbaren Kriterien, auf einer Einkommensumrechnung. In Tabelle 6 ist nur ein erster Schritt dieser Transformation uber eine Umrechnung der Werte aus der jeweiligen Landeswahrung in eine gemeinsame Wahrung versucht worden. Dadurch werden die Betrage in eine Groenordnung gegeben. Eine Vergleichbarkeit bedeutet dieses aber noch nicht, denn, vorausgesetzt die Haushaltsstruktur liee sich uber Informationen zu den Personen im Haushalt auf eine vergleichbare Ebene bringen, so blieben doch die genannten Betrage fur die USA und Grobritannien zunachst auf einen vergleichbaren Zeitraum zu transformieren und sodann um den Anteil der Steuern zu reduzieren. In einem weiteren Schritt mu fur alle vier zu vergleichenden Lander berucksichtigt werden, was man mit einer Summe X im jeweiligen Land an Kaufkraft besitzt. Damit bleibt die Frage: Welches Einkommen hat der konkrete Haushalt wirklich zur Verfugung fur laufende Verpflichtungen wie Bildung und Ausbildung (in den genannten vier Landern finanziell auf sehr unterschiedlichem Niveau), Kommunikation,

Transport, etc. sowie für das tägliche Leben und den Konsum? Die Ermittlung der nationalen Kaufkraft erfordert damit mindestens eine weitere Umrechnung.

Diese zweite Stufe der Transformation muß die unterschiedlichen nationalen Preisniveaus berücksichtigen: Im Land A hat z.B. 1 ECU eine andere Kaufkraft wie in Land B. Die Lösung zum Umgehen der Probleme Devisenkurse und nationaler Preisniveaus ist die Überführung der nationalen Währungen bzw. der Referenzwährung in Kaufkraftindizes und die anschließende Festlegung der Kaufkraftparitäten (Purchasing Power Parity). Der Kaufkraftindex basiert auf den Preisen bestimmter Güter und Dienstleistungen eines Warenkorbes. Aber welche Waren in die Konstruktion des Index eingehen und wie die Gewichtung der einzelnen Preise durchgeführt wird, ist trotz Vorschlägen zur Vereinheitlichung (z.B. durch das International Comparison Project des UN Statistical Office) nationenabhängig. Dennoch stellen die Kaufkraftparitäten ein Maß für einen Geldbetrag dar, für den in einem spezifischen Jahr die gleichen Dienstleistungen, Waren und Güter in den verschiedenen Ländern erworben werden können (World Product and Income 1982). Eine Erweiterung, die auch die unterschiedlichen nationalen Besteuerungssysteme berücksichtigt, ist der Purchasing Power Standard (Eurostat 1983). Da für „Kaufkraftparitäten“ die Qualitätskriterien der Indikatorenkonstruktion gelten, können Meßfehler kontrolliert und damit Reliabilität und Validität von Einkommenserfassungen in verschiedenen Staaten erhöht werden. Nur für nichtmonetäre Einkommensstypen (z.B. Sachleistungen) und nichtquantifizierbare Leistungen und Güter (z.B. Hausarbeit oder Produkte aus der eigenen Landwirtschaft) ist die Kaufkraftparität kein geeigneter Umrechnungsfaktor (vgl.: Zaidi 1991).

Um unterschiedliche Einkommensverteilungen miteinander vergleichen zu können, haben sich der Konzentrationsindex von Gini und daraus abgeleitete Maßzahlen bewährt. Der Gini-Index basiert auf der Lorenzkurve. Der Graph der Lorenzkurve trägt auf der Abszisse die aufaddierten Prozentanteile der Einheiten ab, deren Einkommen betrachtet werden soll. Die Ordinate ist deren prozentualer Anteil am summierten Gesamteinkommen eines Staates. Die Einkommensanteile sind nach ihrer Höhe geordnet (zur Lorenzkurve und den darauf basierenden Indizes siehe: Cowell 1977: 22 ff.; Sen 1975: 42 ff.; vgl. auch: Atkinson/Rainwater/Smeeding 1995).

Um Einkommen bei unterschiedlichen Haushaltsgrößen, Haushaltszusammensetzungen und dem unterschiedlichen Bedarf von Haushaltsmitgliedern unterschiedlicher Generationen und Positionen vergleichbar zu machen, wurden sogenannte Äquivalenzskalen entwickelt. Das mit Hilfe dieser Skalen umgerechnete Einkommen eines Haushaltes wird als Wohlstandsniveau interpretiert. Es stellt ein Maß für den Wohlstand einer

Person in unterschiedlichen Untersuchungseinheiten (z.B. Haushalt) dar (vgl.: Buhmann et al. 1988; Klein 1986).

5. Exkurs: Zur Definition von Haushalt

Schon bei unterschiedlichen nationalen Studien wird mit unterschiedlichen Haushaltsdefinitionen gearbeitet: So definiert die amtliche Statistik „Haushalt“ anders als die Umfrageforschung (siehe unten). Hinzu kommt eine subjektive Anwendung der Definition von „Haushalt“ durch die Interviewer und (eine eventuell hiervon abweichende) durch die Befragungspersonen. Begibt man sich auf die internationale Ebene, so wird die Diskrepanz zwischen den objektiven Definitionen größer und eine Einordnung der subjektiven Sichtweisen der vorgegebenen Definitionen ist nicht mehr möglich. Grundsätzlich kann „Haushalt“ auf vier unterschiedliche Arten definiert werden:

- a) als Wirtschaftseinheit
- b) als Wohneinheit
- c) als Familie
- d) als Netzwerk

Wirtschaftseinheit, Wohneinheit, Familie und Netzwerk erschließen unterschiedliche Perspektiven der Untersuchung. Die Beobachtungseinheiten unterscheiden sich in ihrer Struktur, in ihrer Zusammensetzung, in ihrer Funktion und Leistung, in ihrem Verhalten, in ihrer Stabilität und Dynamik und in ihren sozialen Beziehungen untereinander. Schwierigkeiten der Beobachtung ergeben sich zusätzlich dadurch, daß neben einer sozialwissenschaftlichen Abgrenzung der Untersuchungseinheit, ökonomische Grenzziehungen (Konsumgemeinschaft, Einkommensgemeinschaft), legale Setzungen (Bedarfsgemeinschaft im Sinne eines Sozialhilfegesetzes, Steuerpflicht im Sinne eines Einkommenssteuergesetzes), räumliche Bestimmungen (Wohnung) und gemeinsam geteilte Interessen der Mitglieder (z.B. Wohngemeinschaft) bestehen.

Der „Haushalt“ als Wirtschaftseinheit gründet auf gemeinsamen Entscheidungen und/oder der gemeinsamen Verfügungsgewalt über Einnahmen und Ausgaben. „Haushalt“ als Wirtschaftseinheit wird im internationalen Kontext aber unterschiedlich operationalisiert. In der Bundesrepublik Deutschland wird die Wirtschaftseinheit „Haushalt“ in der Regel über den gemeinsamen Herd definiert: „zusammen leben und wirtschaften“ (Mi- krozensus, EVS). Dieses kann allerdings auf eine (familiäre) Hausgemeinschaft über mehr als eine Wohnung erweitert werden zu „gemeinsam wohnen und wirtschaften, d.h. ihren Lebensunterhalt gemeinsam finanzieren“ (deutsche Haushaltsdefinition im ECHP, StaBA 1997a: Pkt. 5.2). Im Gegensatz hierzu werden in den Niederlanden, aber auch in vielen anderen Kulturen Südeuropas als Kriterium der Definition die gemeinsamen Mahlzeiten gesehen: Personen in einer Wohnung, die gemeinsam Mahlzeiten einnehmen. In Frankreich wird die Wirtschaftseinheit „Haushalt“ entweder, wie in der Bundesrepublik Deutschland, über den gemeinsamen Herd oder über das gemeinsame Haushalten definiert: Personen die gemeinsam zur Steuer veranlagt sind. In Großbritannien werden drei Elemente der Definition kombiniert: Personen, die eine gemeinsame Adresse haben, gemeinsam Mahlzeiten zubereiten und eine gemeinsame Haushaltsführung haben.

„Haushalt“ als Wohneinheit definiert sich über „(private) dwelling“, den Wohnraum, die Wohnadresse. Aber auch hier wird unterschiedlich operationalisiert. In der Bundesrepublik werden unter dieser Definition Personen verstanden, die in derselben Wohnung zusammenleben. In Australien sind es diejenigen, die sich eine gemeinsame Küche

teilen. Die gemeinsame Küche kommt dem gemeinsamem Herd schon sehr nahe. Demgegenüber ist in Luxemburg das gemeinsame Wohnzimmer ausschlaggebend.

Die Definition des „European Community Household Panel“ (ECHP) vereint eine Vielzahl der Elemente der auf Wirtschaftseinheit oder auf Wohneinheit basierenden nationalen Definitionen: „... a household is defined ... in terms of two criteria: the sharing of the same dwelling, and the common living arrangements. ... The shared arrangements may include meals taken together or a shared room ... and/or a joint budget ... and/or the use of common equipments“ (Europäische Kommission 1996: 17). Allerdings stellt sich die Frage, ob ein so definierter Haushalt, der die Definitionen fast aller an der Umfrage beteiligten Nationen bzw. Kulturen berücksichtigt, im Endergebnis nicht doch wieder das Problem der Vergleichbarkeit aufweist, denn jede Nation wird den ihrer Kultur am nächsten kommenden definitorischen Merkmalen den Vorrang geben. Damit liegt zwar eine Definition vor, die sich jedoch sehr unterschiedlich interpretieren läßt und damit Definitionsunterschiede hervorhebt (siehe StaBA 1997a: Pkt. 5.2).

Die dritte Art der Definition von „Haushalt“ gründet auf dem Begriff der „Familie“. Familie definiert sich über das Zusammenleben von mindestens zwei Personen, die über Geburt, Heirat oder Adoption in Beziehung zueinander stehen. Ein Familienhaushalt in diesem Sinne ist ein Haushalt, den durch Geburt, Heirat oder Adoption miteinander verbundene und in einer „housing unit“ zusammenlebende Personen unterhalten. Zusätzliche, nicht von der Definition der Familie erfaßte Personen, können weitere Mitglieder im Familienhaushalt sein. Diese Haushaltsdefinition ist die in den USA vorherrschende (vgl.: U.S. Bureau of the Census 1996, S. A-1f.). Eine zweite Möglichkeit der Definition des Familienhaushalts ist das dauerhafte Zusammenleben von Personen in einer Wohneinheit. Das konstituierende Element dieser Definition ist ein potentielles Abhängigkeitsverhältnis der Mitglieder untereinander (Curtis 1986).

Die vierte Art der Definition von „Haushalt“ baut auf einem engen Begriff sozialer Hilfsnetzwerke auf und „beschwört“ entweder die „haushaltstypische“ gegenseitige Hilfe oder beinhaltet „processes and relationships that compose the systems of reproduction of labor power“ (Smith/Wallerstein/Evers 1984: 8). Räumliche Operationalisierungskriterien der Einheiten (wie Wohnung, Küche) verlieren bei diesem Ansatz an Bedeutung. Lediglich die räumliche Bedingtheit der Netzwerkbeziehungen bleiben erhalten in Form von Erreichbarkeit für Kommunikation und Interaktion im Netz. „We have thus reconceptualized households not as a determinate set of people but as a set of relationships that impose a mutual obligation to pool resources from a multiplicity of labor forms whether or not one of those resources is a common residence“ (Friedman 1984: 48). Angesichts demographischer und kultureller Veränderungen der

modernen Gesellschaften, bei der klassische Formen des „Haushaltes“ und der „Familie“ durch neue Formen der Lebensführung abgelöst werden, gewinnt der Austausch „haushaltstypischer“ Ressourcen und Leistungen im sozialen Netzwerk immer mehr an Bedeutung. Der am Studienort lebende Studierende kann trotz räumlicher Trennung auf die Funktionen seines Ursprungshaushaltes zurückgreifen; moderne Kommunikations- und Transporttechniken ermöglichen die Trennung vom Ort des Haushaltes und vom Ort des Einkommenserwerbs.

6. Probleme der Vergleichbarkeit

Jede Abfrage von Haushaltseinkommen birgt eine Reihe von Fehlerquellen in sich, die eine Vergleichbarkeit mit anderen Studien national oder international erschweren. Einige der jetzt anzuführenden Punkte beziehen sich allgemein auf die Umfrageforschung und die Durchführung von Umfragen, andere betreffen spezifisch die Einkommenserfassung. Zu beachten sind alle hier genannten Punkte gerade bei vergleichenden Umfragen bzw. beim Heranziehen von Umfragedaten unterschiedlicher Erhebungen für nationale wie internationale Vergleiche:

Problem Nr. 1: Die Stichprobe und die Erreichbarkeit der Zielperson

Wichtig für die Vergleichbarkeit von Studien ist der Stichprobenplan: Sind die Stichproben der zum Vergleich anstehenden Studien nach vergleichbaren Regeln gezogen worden? Handelt es sich um Zufallsstichproben? Wird die Zielperson über eine Zufallsauswahl aus einem Register oder über eine Zufallsauswahl aus den Mitgliedern kontaktierter Haushalte ermittelt? Und wie werden diese Haushalte ausgewählt? Geschieht dieses über eine Begehung (nach welchen Regeln?) oder über ein Register? (Siehe auch Gabler/Hoffmeyer-Zlotnik 1997.)

Die Stichprobenziehung für das ECHP z.B. basiert auf unterschiedlichen nationalen Stichprobenplänen und Grundgesamtheiten: den Bevölkerungsregistern in Luxemburg und Italien, dem Verzeichnis der Postadressen in Großbritannien und dem Census in Spanien. In der Bundesrepublik Deutschland ist es die Wohnbevölkerung, aus der über eine Begehung nach Random-Route aufgelistet wird. Die Stichprobenziehung geschieht (außer in Luxemburg) zweistufig und hat mindestens zwei zufällig ausgewählte „sampling points“ pro Schicht. Zwar sahen die von Eurostat vorgegebenen Regeln eine einheitliche Stichprobe für alle Teilnehmerländer am ECHP vor - jedoch wurden nationale Stichproben nach nationalen Stichprobendesigns realisiert, deren Konzepte nicht immer mit den Eurostat-Vorgaben in Übereinklang stehen. Dieses Dilemma begleitet eine Vielzahl von international vergleichenden Studien.

Nur eine Zufallsauswahl von Haushalten oder Personen, die sich mit geringer lokaler Klumpung über eine hohe Anzahl von „sampling points“ verteilen, garantieren ein akzeptables Abbild der vorhandenen Haushaltstypen. Allerdings haben bei Personenstichproben, die aus Personenregistern gezogen werden, kleine Haushalte eine geringere Chance in die Stichprobe zu gelangen (jede Person mit dem Gewicht 1 wird multipliziert mit der Anzahl der zur Grundgesamtheit zählenden Haushaltsmitglieder), als dieses bei der Zielpersonenermittlung über den Umweg einer Haushaltsstichprobe möglich ist (jeder Haushalt mit dem Gewicht 1 wird dividiert durch die Anzahl der zur Grundgesamtheit zählenden Haushaltsmitglieder).

Sind die zu vergleichenden Studien nach vergleichbaren Stichprobenplänen durchgeführt worden, so folgt der Versuch der Klassifikation des Non-Response: Es gibt eine unterschiedlich gute Erreichbarkeit unterschiedlicher Haushaltstypen und es gibt gruppenspezifische Verweigerer der Teilnahme von Umfragen sowie der Beantwortung heikler Fragen.

Problem Nr. 2: Die Definition von Haushalt

Die Definition von „Haushalt“ geschieht beim Forscher nach kulturellen und strukturellen Merkmalen. Damit ist der Haushaltsbegriff auf der Seite des Forschers einerseits abhängig von der Forschungsfrage und andererseits abhängig von den kulturellen Vorgaben. Die subjektive Haushaltsdefinition bei den Interviewern und/oder bei den Zielpersonen basieren auf deren jeweiligem Selbstverständnis und auf der Praktikabilität der Definition bei der Ermittlung der nachgefragten Informationen und kann damit von der Haushaltsdefinition des Forschers abweichen.

Da der Haushaltsbegriff von einer eng definierten kleinen Wirtschaftseinheit bis zur wesentlich weiter definierten (Groß-)Familie und zum Netzwerk reichen kann, ergibt die Berechnung von Haushaltseinkommen auf der Basis unterschiedlicher Definitionen auch sehr unterschiedliche Werte, die sich nicht durch Währungs- oder Kaufkraftgewichtungen wieder ausräumen lassen. Eine Vergleichbarkeit von Haushalt bedingt die Anwendung einer Haushaltsdefinition, die unter kulturellen Bedingungen von allen in den Vergleich einbezogenen Nationen/Kulturen akzeptiert werden kann, vorausgesetzt man hat die hierzu notwendigen Daten zur Analyse von Struktur und Zusammensetzung der Haushalte im benötigten Umfang erfaßt.

Problem Nr. 3: Die Operationalisierung der Einkommensabfrage

Wie aufgezeigt wurde, wird Einkommen national wie international über die einzelnen Studien gemäß der Forschungsfrage unterschiedlich operationalisiert. Unterschiedliche Operationalisierungen generieren unterschiedliche Antwortstrukturen. Bei der Operationalisierung von vergleichenden Einkommensabfragen kommt es nicht nur darauf an, daß

für die befragte Person die Definition von Einkommen eindeutig ist, sondern daß die Vorgaben auch hinsichtlich der Berechnung der Summe, des zeitlichen Bezugs und des personellen Bezugs vergleichbar sind.

Problem Nr. 4: Die cognitive Leistung des Erinnerns und Addierens von „Einkommen“

Ein zentrales Problem stellen die cognitiven Leistungen der Befragten dar: Was wird erinnert? In der Regel wird sich die Berechnung des Haushaltseinkommens über die regelmäßig fließenden Arbeitseinkommen und Altersrenten, -pensionen definieren. Öffentliche Transfers werden ohne spezielle Erwähnung durch das Fragenprogramm nur insofern erinnert, soweit diese den überwiegenden Lebensunterhalt darstellen. Private Transfers werden in der Regel nicht erinnert, da sie einerseits als selbstverständlich und andererseits als nicht zum Einkommen zählend betrachtet werden. Einmalige öffentliche Transfers (z.B. der Lohnsteuerjahresausgleich) haben bei der Abfrage weder in der Erinnerung noch im Ergebnis einen großen Stellenwert. Vermögen betrifft die Masse der Haushalte nicht - wird in der Regel aber dort, wo es ins Gewicht fallen sollte, ohne explizites Nachfassen nicht benannt. Deputate bleiben, oft unbeabsichtigt, eine black box: Weder der Dienstwagen, noch die Werkswohnung werden als zusätzliches Einkommen betrachtet.

Problem Nr. 5: Der Kenntnisstand des Informanten

Bei den meisten allgemeinen Bevölkerungsumfragen der Markt- und Sozialforschung wird eine Referenzperson die Einkommensfrage stellvertretend für den ganzen Haushalt beantworten. Dieses setzt voraus, daß die Referenzperson als Experte zur Einkommenssituation aller Haushaltsmitglieder anzusehen ist und als Experte gut informiert ist. Über das Haushaltseinkommen sind in der Regel aber nur jene Personen relativ gut informiert, die nahe dem Zentrum des Haushaltsgeschehens stehen und zum ökonomischen Zentrum der Wirtschaftseinheit „Haushalt“ gehören. Peripher stehende Personen sind als Auskunftspersonen über den Haushalt und dessen wirtschaftliche Ressourcen in der Regel ungeeignet, haben aber in größeren Haushalten eine gute Chance zur Befragungsperson zu werden. Selbst in Zweipersonenhaushalten mit zwei Befragungspersonen im erwerbsfähigen Alter liegen die Angaben der nicht-erwerbstätigen Befragungspersonen zur Höhe des Haushaltseinkommens um etwa 30 Prozent niedriger als bei den erwerbstätigen Befragungspersonen (siehe: ALLBUS 1996; vgl.: Sozialwissenschaften-Bus 3/96).

Problem Nr. 6: Umrechnung von nationalem Einkommen in international vergleichbare Einheiten

Um auf nationaler Ebene den unterschiedlichen Fehlern bei der Datenerhebung zu entkommen, kann man die Klassifikation der Einkommen über eine Gewichtung am Census oder Mikrozensus orientieren. Dies ist eine Möglichkeit, Verzerrungen und Ausfälle zu klassifizieren, allerdings nicht zu korrigieren. Im internationalen Vergleich fehlen Orientierungsgrößen, um Fehler zu klassifizieren, da weder die Einkommensberechnungen (welche Abzüge vom Brutto-Einkommen werden in welchen Ländern normalerweise

vorgenommen?) noch die Definition der Haushalte (wer zählt zum Haushalt dazu?) kulturunabhängig sind. Um eine Umrechnungsmöglichkeit für den internationalen Vergleich zu erhalten, müssen kulturabhängige nationale Eigenheiten, sowie die Klassifikation der Ausfälle dokumentiert werden (siehe Hackauf et al. 1990). Nur über eine Dokumentation nationaler Eigenheiten lassen sich diese dann zur Interpretation der gefundenen Ergebnisse im Vergleich heranziehen.

Daneben gibt es allerdings Möglichkeiten, Einkommensverteilungen in verschiedenen Ländern mit unterschiedlichen Währungen vergleichbar zu machen in einem ersten Transformationsschritt über das Umrechnen aller im Vergleich auftretenden Währungen in eine Referenz- oder Leit-Währung. In einem zweiten Transformationsschritt können Kaufkraftindizes und Wohlstandsniveaus berechnet und zur Vergleichbarkeit mit herangezogen werden.

Anmerkung

*) Uwe Warner ist wissenschaftlicher Mitarbeiter am Centre d'Etudes de Populations, de Pauvreté et de Politiques Socio-Economiques (CEPS), Rue Emile Mark, L-4501 Differdange, Luxembourg.

Literatur

Atkinson, A.B./L. Rainwater/T.M. Smeeding, 1995: Income Distribution in OECD Countries. Evidence from the Luxembourg Income Study. Paris: OECD.

Buhmann, B./L. Rainwater/G. Schmaus/T.M. Smeeding, 1988: Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates across ten Countries using the Luxembourg Income Study (LIS) Database. *The Review of Income and Wealth*, Vol. 34/2: 115-142.

Cowell, F.A., 1977: *Measuring Inequality*. Oxford: Philip Allan.

Curtis, R.F., 1986: Household and Family in Theory on Inequality. *American Sociological Review*, Vol. 51: 168-183.

Deutsche Forschungsgemeinschaft, 1994: *Mikroanalytische Grundlagen der Gesellschaftspolitik. Sonderforschungsbereiche*. Berlin: Akademie Verlag, 2 Bände.

Europäische Kommission, 1996: *European Community Household Panel (ECHP): Volume 1 - Survey Methodology and Implementation. - Survey Questionnaires*. Luxembourg.

Eurostat, 1983: *Comparison in Real Values of Aggregates of ESA*. Luxembourg.

- Eurostat, 1995: ESSPROS Manual 1995. Vol. 1. General Principles and the Core System. Luxemburg.
- Friedman, K., 1984: Households as Income-Pooling Units. S. 37ff. in: J. Smith/I. Wallerstein/H.-D. Evers (Hrsg.), *Households and the World-Economy*. Beverly Hills: Sage.
- Gabler, S./J. H.P. Hoffmeyer-Zlotnik (Hrsg.) 1997: *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Gräbe, S. (Hrsg.) 1991: *Der private Haushalt als Wirtschaftsfaktor*. Frankfurt/M.: Campus.
- Hackauf, H./R. Randell/J. Coder/T.M. Smeeding/U. Warner, 1990: *Final Report for the Statistical Offices of the European Community Project on Technical Documentation and Income Data Quality for the Luxembourg Income Study Database*. Luxemburg.
- Harkness, J./Mohler, P. Ph./Thomas R. (eds.): 1997: *General Report on Study Programme for Quantitative Research (SPQR)*. Abschlußbericht an die Europäische Kommission für das European Consortium for comparative Social Surveys (COMPASS). Mannheim: ZUMA.
- Harder, B.L. von, 1994: *Die Einkommensermittlung in der ZUMA/GFM-GETAS-Standardstatistik im Methodenvergleich*. Wissenschaftliches Symposium anlässlich der 20-Jahr-Feier des ZUMA. Mannheim: Materialien zum Vortrag.
- Klein, T., 1986: *Äquivalenzskalen - Ein Literatursurvey*. Sfb 3 Projekt C-1: Soziale Sicherung. Arbeitspapier Nr. 195. Frankfurt/M.
- Sen, A., 1975: *Ökonomische Ungleichheit*. Frankfurt/M.: Campus.
- Schmaus, G., o.A.J.: *The PACO-Project User Manual*. Differdange: CEPS.
- Smith, J./I. Wallerstein/H.-D. Evers, 1984: *Households and the World-Economy*. Beverly Hills: Sage.
- Statistisches Bundesamt (Hrsg.), 1984: *Statistisches Informationssystem des Bundes. Definitionskatalog*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (Hrsg.), 1995: *Demographische Standards. Eine gemeinsame Empfehlung des Arbeitskreises Deutscher Marktforschungsinstitute (ADM), der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute (ASI) und des Statistischen Bundesamtes*. Ausgabe 1995. *Methoden-Verfahren-Entwicklungen. Materialien und Berichte*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (Hrsg.), 1997: *Statistisches Jahrbuch 1997 für die Bundesrepublik Deutschland*. Stuttgart: Metzler + Poeschel.

Statistisches Bundesamt (Hrsg.), 1997a: Handbuch für die Benutzung der faktisch anonymisierten Daten der deutschen Stichprobe des Europäischen Haushaltspanels. Wiesbaden: Statistisches Bundesamt.

U.S. Bureau of the Census, 1996: Current Population Reports. Series P-60, No. 193. Money Income in the United States: 1995 (With Seperate Data on Valuation of Noncash Benefits). Washington D.C.: U.S. Government Printing Office.

Warner, U., 1994: The Luxembourg Income Study as an Instrument for Comparative Sociological, Socio-Economic, and Socio-Political Research. in: P. Flora/F. Kraus/H.-H. Noll/ F. Rothenbacher (Hrsg.), Social Statistics and Social Reporting in and for Europe. Bonn: Informationszentrum Sozialwissenschaften.

World Product and Income: International Comparison of Real Gross Product, 1982. Baltimore: John Hopkins University Press.

Zaidi, M.A., 1991: Purchasing Power Parity. Discussion paper EURDP91.07. Department of Economic Sociology and Psychology, Erasmus University. Rotterdam.

Materialien zur Durchführung dieser Studie

Um diesen Aufsatz zu realisieren, haben wir auf Dokumente, Handbücher und Fragebögen von nationalen und internationalen Umfragen zurückgegriffen. Auch die Hilfe von Kolleginnen und Kollegen, die in den jeweiligen Ländern als Experten für nationale und vergleichende Umfragen arbeiten, war von unschätzbarem Wert. Die benutzten Quellen sind: Für die USA: die Fragebögen zu den Umfragen: „March Current Population Survey“ (im Anhang von U.S. Bureau of the Census 1996 abgedruckt), „Panel Study of Income Dynamics“ und „General Social Survey“ vom NORC in Chicago. Für die *Bundesrepublik Deutschland*: Erhebungsinstrumente der amtlichen Statistik: „Einkommens- und Verbrauchsstichprobe 1993“ und „Mikrozensus 1995“; Erhebungsinstrumente der akademischen Sozialforschung: „Wohlfahrtssurvey 1993“ (WZB, ZUMA), „Sozio-ökonomisches Panel - SOEP 1993“ (DIW) und „Allgemeine Bevölkerungsumfrage in den Sozialwissenschaften - ALLBUS 1996“ (ZUMA). Des weiteren wurden die Fragebögen der GFM-GETAS-Großumfrage 1993, des Sozialwissenschaften-Bus 3/93 (ZUMA-Standarddemographie) und der Media-Analyse 1993 mit einbezogen, sowie die „Demographischen Standards“, herausgegeben vom Arbeitskreis Deutscher Marktforschungsinstitute, der Arbeitgemeinschaft Sozialwissenschaftlicher Institute und dem Statistischen Bundesamt (die neue Auflage erscheint im Frühsommer 1998), betrachtet. Das *luxemburger* Beispiel wurde vom Institut Luxembourgeois de Recherche Sociales et d'Etudes de Marche (ILRES) zur Verfügung gestellt. *International vergleichend* dienten uns Materialien und Daten des „International Social Survey Programme - ISSP“ und der daran beteiligten Institute, des „European Community Household Panel“ und der daran beteiligten nationalen statistischen Ämter. Als weiteres für den internationalen Einsatz

konzipiertes Instrument wurden Fragebogen und Ausführungen der „Working Party on Harmonisation of Demographics“ der ESOMAR betrachtet. Hilfreich waren auch die Arbeiten zum Projekt „Study Programme for Quantitative Research“ (Harkness, J./Mohler, P. Ph./Thomas R. 1997).

Anhang: Die Operationalisierung von Einkommen in Studien

Die Darstellung der Fragen zur Erfassung des Einkommens erfolgt für alle Studien nach folgendem Schema:

- Selbständig erwerbstätig versus nicht selbständig erwerbstätig
- Bruttobetrag versus Nettobetrag
- Zeitlicher Bezug: Monat versus Jahr
- Markteinkommen (dieses sind Einkommen aus Erwerbstätigkeit und über Deputate sowie Einkommen aus Besitz und Vermögen) und/oder Transfereinkommen (Einnahmen aus Einkommensübertragungen)
- Individualeinkommen versus Haushaltseinkommen
- Offene versus kategoriale Abfrage
- Gemessen versus geschätzt/objektive versus subjektive Beträge
- Nur monetäre oder auch nicht-monetäre Einkommensarten
- „Ausführliche“ oder „kaum vorhandene“ Erläuterungen/Definition von Einkommen
- Umfang der Abfrage/Anzahl der Fragen

Abfragen von Einkommen in der amtlichen Statistik

Mikrozensus

- selbständig Erwerbstätige: über „Unternehmereinkommen“ erfaßt
- Nettoeinkommen:
 - für nicht-selbständig Erwerbstätige: Einkommen abzüglich „Lohnsteuer“, „Kirchensteuer“, „Sozialversicherung u.ä. Beiträge“,
 - für selbständig Erwerbstätige: ohne Abzüge nicht klar definiert
- zeitlicher Bezug: Stichmonat
- Definition von Markteinkommen: Lohn, Gehalt, Unternehmereinkommen, Renten (sehr differenziert), Pensionen, Vermietung, Verpachtung, Gratifikation, vermögenswirksames Sparen, Vorschüsse, Werkswohnungsmiete
- Definition von Transfereinkommen: sehr differenzierte Auflistung öffentlicher und privater Transferleistungen
- Fokussierung auf Einkommenstyp: über „überwiegenden Lebensunterhalt“
- Individualeinkommen vs. Haushaltseinkommen: Haushaltseinkommen als Aggregat aller Individualeinkommen der Mitglieder eines Haushalts
- offene Abfrage: nein
- kategoriale Abfrage: in 18 Kategorien, Schrittweite: 300/400/500 bis DM 7500,-
- gemessen vs. geschätzt (durch wen?): Befragtenantwort ohne schätzen
- monetär vs. nicht-monetär: es werden Einnahmen monetär in DM erhoben
- Erläuterung/Definition von Einkommen: sehr ausführlich in der Abfrage (Kategorien, Anweisungen) sowie in separaten Erläuterungen

- Umfang der Abfrage/Anzahl der Fragen: eine Einkommensfrage je Individuum, die im Zusammenhang mit der Abfragen der Einkommensstypen steht.

Einkommens- und Verbrauchsstichprobe - EVS

- Selbstständige: über „Soziale Stellung“
- Nettoeinkommen:
 - für nicht-selbstständig Erwerbstätige: abzüglich Steuern und Pflichtbeiträgen zur Sozialversicherung,
 - für selbstständig Erwerbstätige: letzte Steuererklärung
- zeitlicher Bezug: monatlich und vierteljährlich
- Definition von Markteinkommen: Einkommen aus Erwerbstätigkeit, Renten, Pensionen (auf Individualebene: Renten der gesetzlichen Rentenversicherung, u.ä.), Vorruhestand, Vermietung, Verpachtung, Geldvermögen, Einkommen aus Gewerbebetrieb oder selbständiger Tätigkeit; als monatlicher Durchschnitt: 13./14. Monatsgehalt, Urlaubsgeld, Dividenden
- Definition von öffentlichen Transfereinkommen: Wohngeld, Kindergeld, Mutterschaftsgeld, Erziehungsgeld, BAFöG, Arbeitslosengeld, Arbeitslosenhilfe, laufende und einmalige Sozialhilfe, Kurzarbeits-, Schlechtwettergeld, Renten der gesetzlichen Unfallversorgung, Renten der Kriegsopferversorgung, sonstige Zahlungen der öffentlichen Kassen.
- Definition von privaten Transfers: Unterhaltszahlungen, Unterstützung durch Verwandte o.ä.
- Fokussierung auf „überwiegender Lebensunterhalt“
- Haushaltseinkommen als Summe der monatlichen Nettoeinkommen aller Haushaltsmitglieder
- offene Abfrage: nein
- kategoriale Abfrage: in 32 Kategorien, Schrittweite: in 200/300 Schritten bis DM 5.000,-, dann in 500er Schritten bis DM 10.000,-, dann in 5.000er bis DM 35.000,-
- gemessen vs. geschätzt (durch wen?): gemessen durch Befragtenantwort
- monetär vs. nicht-monetär: es werden Einnahmen und Ausgaben monetär in DM erhoben
- Erläuterung/Definition von Einkommen: ausführliche Definitionen und Abfragelisten
- Umfang der Abfrage/Anzahl der Fragen: je Individuum 55 Items im Tagebuch über ein Viertel Jahr

European Community Household Panel ECHP

- selbständig Erwerbstätige: „self-employment“ (im Sinne von „Eigentümer von ...“) und „farming“;
- Bruttoeinkommen für Individuum:
 - für nicht-selbstständig Erwerbstätige: vor Steuern und Sozialabgaben,
 - für selbstständig Erwerbstätige: „Profit“ nach Abzug aller Ausgaben und Löhne, aber vor Steuern
- Nettoeinkommen für Individuum:
 - für nicht-selbstständig Erwerbstätige: „the take home pay“,
 - für selbstständig Erwerbstätige: nicht definiert

- Nettoeinkommen für Haushalt:
- für nicht-selbständig Erwerbstätige: nach Steuern und Abgaben zur Sozialversicherung und zur Pensionskasse,
- für selbständig Erwerbstätige: nicht definiert
- zeitlicher Bezug:
- Jahr: für die einzelnen Monate des Referenzjahres
- Monat: für den aktuellen Monat, in dem das Interview stattfindet
- Definition von Markteinkommen: detaillierte Auflistung der Einkommenstypen
- Definition von Transfereinkommen: detaillierte Auflistung der Einkommenstypen
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: beide mit separaten Fragen
- offene Abfrage: ja
- kategoriale Abfrage: in 8 Kategorien, Schrittweite 500/1000 ECU
- gemessen vs. geschätzt (durch wen?): gemessen durch Befragtenantwort; Befragter hat auch die Möglichkeit des Schätzens
- monetär vs. nicht-monetär: es werden Einnahmen monetär in Landeswährung erhoben
- Erläuterung/Definition von Einkommen: ausführliche Erläuterungen in der Fragenformulierung und in separaten Interviewerheften
- Umfang der Abfrage/Anzahl der Fragen: 24 Fragen zu Haushaltseinkommen und 29 Fragen zu Individualeinkommen sind insgesamt 53 Fragen, zuzüglich monatliche Abfrage mit je 2 Fragen für 12 Monate

Abfragen von Einkommen in der deutschen Umfrageforschung

ZUMA-Standarddemographie; Demographische Standards

- selbständig Erwerbstätige: über „Einkommen aus selbständiger Tätigkeit“ erfaßt
- Nettoeinkommen:
- für nicht-selbständig Erwerbstätige: abzüglich „Steuern und Sozialversicherungsbeiträge“,
- für selbständig Erwerbstätige: abzüglich „Betriebsausgaben“
- zeitlicher Bezug: Monat
- Definition von Markteinkommen: Lohn, Gehalt, Einkommen aus selbständiger Tätigkeit, Rente, Pension, Vermietung, Verpachtung
- Definition von Transfereinkommen: Kindergeld, Wohngeld, Sozialhilfe, „sonstige Einkünfte“
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: beide Einkommen in getrennter Abfrage
- offene Abfrage: ja; bei Verweigerung: kategorial
- kategoriale Abfrage: in 22 Kategorien, Schrittweite: 200/ 250/ 500 bis DM 6000,-
- gemessen vs. geschätzt (durch wen?): Befragtenantwort ohne schätzen
- monetär vs. nicht-monetär: es werden Einnahmen monetär in DM erhoben
- Erläuterung/Definition von Einkommen: ausführliche Definition von Einkommen in Abfrage und Intervieweranweisung

- Umfang der Abfrage/Anzahl der Fragen: 2 zweistufige Fragen (offen; wenn Verweigerung: geschlossen nachfragen)

ALLBUS; International Social Survey Programme (Deutschland)

- selbständig Erwerbstätige: über „Selbständige“ erfaßt
- Nettoeinkommen:
 - für nicht-selbständig Erwerbstätige: abzüglich „Steuern und Sozialversicherungsbeiträge“,
 - für selbständig Erwerbstätige: abzüglich „Betriebsausgaben“
- zeitlicher Bezug: Monat
- Definition von Markteinkommen: nicht definiert
- Definition von Transfereinkommen: nicht definiert
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: beide Einkommen in getrennter Abfrage
- offene Abfrage: ja; bei Verweigerung: kategorial
- kategoriale Abfrage: in 22 Kategorien, Schrittweite: 200/ 250/ 500 bis DM 6000,-
- gemessen vs. geschätzt (durch wen?): Befragtenantwort ohne schätzen
- monetär vs. nicht-monetär: es werden Einnahmen monetär in DM erhoben
- Erläuterung/Definition von Einkommen: unzureichende Definition von Einkommen
- Umfang der Abfrage/Anzahl der Fragen: 2 zweistufige Fragen (offen; wenn Verweigerung: geschlossen nachfragen); 1 Frage subjektive Einschätzung

Media-Analyse - MA

- selbständig Erwerbstätige: kein Bezug
- Nettoeinkommen:
 - für nicht-selbständig Erwerbstätige: abzüglich „Steuern und Sozialversicherung“,
 - für selbständig Erwerbstätige: ohne Abzüge, nicht klar definiert
- zeitlicher Bezug: Monat
- Definition von Markteinkommen: „selbst verdienen“, „Einkommen aus Renten, Pensionen, Mieten, Pacht, Wertpapieren und dergleichen“
- Definition von Transfereinkommen: nicht vorhanden
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: beide Einkommen in getrennter Abfrage
- offene Abfrage: nein
- kategoriale Abfrage: in 12 Kategorien, Schrittweite: 500 bis DM 5000,-
- gemessen vs. geschätzt (durch wen?): Befragtenantwort gemessen oder geschätzt (zweistufige Abfrage) oder Schätzung durch den Interviewer (mit Kennung)
- monetär vs. nicht-monetär: es werden Einnahmen monetär in DM erhoben
- Erläuterung/Definition von Einkommen: sehr knappe Erläuterungen im Fragentext
- Umfang der Abfrage/Anzahl der Fragen: 3 Fragen

Omnibusumfrage (GFM-GETAS Großumfrage)

- selbständig Erwerbstätige: kein Bezug
- Nettoeinkommen:
 - für nicht-selbständig Erwerbstätige: kein Bezug,
 - für selbständig Erwerbstätige: kein Bezug
- zeitlicher Bezug: Monat
- Definition von Markteinkommen: nicht vorhanden
- Definition von Transfereinkommen: nicht vorhanden
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: Haushaltseinkommen
- offene Abfrage: nein
- kategoriale Abfrage: 16 Kategorien, Schrittweite: 250 bis 3000 DM, dann 1000 bis 6000 DM
- gemessen vs. geschätzt (durch wen?): Befragtenantwort
- monetär vs. nicht-monetär: es wird Einkommen monetär in DM erhoben
- Erläuterung/Definition von Einkommen: Reduktion auf zwei Worte: „Monatliches Haushalts-Netto-Einkommen“
- Umfang der Abfrage/Anzahl der Fragen: Reduktion auf eine Liste mit Überschrift

Sozio-ökonomisches Panel - SOEP

- selbständig Erwerbstätige: definiert über „Gewerbebetrieb“ oder „selbständige Tätigkeit“
- Bruttoeinkommen:
 - für selbständig Erwerbstätige: monatliche Einkünfte aus Gewerbebetrieb oder selbständiger Tätigkeit laut Einkommensteuerbescheid,
 - für alle: Transferzahlungen
- Nettoeinkommen:
 - für nicht-selbständig Erwerbstätige: Arbeitseinkommen nach Abzug von Steuern und Sozialabgaben zuzüglich Wohngeld, Kindergeld, BAföG und Unterhaltszahlungen
- zeitlicher Bezug: Monat
- Definition von Markteinkommen: Arbeitsverdienst minus Lohnsteuer minus Sozialbeiträge
- Definition von Transfereinkommen: wird detailliert und vollständig nach Liste abgefragt (brutto pro Monat)
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: es werden Individual- und Haushaltseinkommen separat erfaßt:
 - Individualeinkommen: separat erfragt: brutto/netto je Einkommenstyp; erfaßt werden Arbeitseinkommen plus Listen von Einkommenstypen jeweils über Hauptliste für Typ und ergänzender spezifischer Nebenliste
 - Haushaltseinkommen: in netto für spezifische Einkommenstypen; zusammenfassend
- offene Abfrage: ja
- kategoriale Abfrage: nein

- gemessen vs. geschätzt (durch wen?): Befragtenantwort gemessen und notfalls geschätzt
- monetär vs. nicht-monetär: es werden Einnahmen monetär in DM erhoben
- Erläuterung/Definition von Einkommensart: ausführliche Erläuterungen in der Fragenformulierung zuzüglich ausführlicher Definitionen der Begriffe und Einkommensarten in einem Interviewerheft
- Umfang der Abfrage/Anzahl der Fragen: 33 Fragen

Abfragen von Einkommen in der internationalen Umfrageforschung

ESOMAR, Eurobarometer:

- selbständig Erwerbstätige: Gruppe wird im Vorfeld definiert
- Einkommen: wird nicht erhoben
- zeitlicher Bezug: Zeitpunkt Befragung
- Definition von langlebigen Konsumgütern statt Markt- und/oder Transfereinkommen
- Individualeinkommen vs. Haushaltseinkommen: „besitzen Sie oder ein anderes Mitglied des Haushalts ...“
- offene Abfrage: nein
- kategoriale Abfrage: in 10 Kategorien, ja/nein-Antworten
- gemessen vs. geschätzt (durch wen?): Befragtenantwort
- monetär vs. nicht-monetär: nicht-monetär über langlebige Konsumgüter und Hausbesitz (des Hauptwohnsitz)
- Erläuterung/Definition von Einkommen: keine
- Umfang der Abfrage/Anzahl der Fragen: 2 Fragen

International Social Survey Programme - ISSP, international:

- selbständig Erwerbstätige:
- Nettoeinkommen: Österreich (A), Italien (I)
- für nicht-selbständig Erwerbstätige: A: nicht definiert, I: „global net income“,
- für selbständig Erwerbstätige: nicht definiert
- Bruttoeinkommen: Großbritannien (GB) und Nordirland (NIRL), die USA, Ungarn (H), die Niederlande (NL), Republik Irland (IRL)
- für nicht-selbständig Erwerbstätige: nicht definiert
- für selbständig Erwerbstätige: nicht definiert
- zeitlicher Bezug:
- Jahr: USA: „last year“, Rep. Irland: „annually“
- Monat: GB, NIRL, A, H, NL, I
- Definition von (Markt)Einkommen: GB, NIRL: „all sources“, IRL: „average household income“, NL: „household income“, USA: „total family income“, A: „all family members income“, I: „global net income“
- Definition von Transfereinkommen: H: „incl. social benefits“, andere: keine Definition
- Fokussierung auf Einkommensstyp: keine

- Individualeinkommen vs. Haushaltseinkommen: beide Einkommen in getrennter Abfrage
- offene Abfrage: für einige Länder
- kategoriale Abfrage: bis zu 22 Kategorien, Schrittweite: unterschiedlich nach Ländern
- gemessen vs. geschätzt (durch wen?): Befragtenantwort
- monetär vs. nicht-monetär: es werden Einnahmen monetär in Landeswährung erhoben
- Erläuterung/Definition von Einkommen: zumeist sehr knapp
- Umfang der Abfrage/Anzahl der Fragen: 2 Fragen, teils zweistufig offen und kategorial

General Social Survey - GSS, USA:

- selbständig Erwerbstätige: „earned any money from any job or employment“
- Bruttoeinkommen: „before taxes“
- für nicht-selbständig Erwerbstätige: keine Definition
- für selbständig Erwerbstätige: keine Definition
- zeitlicher Bezug: „last year“
- Definition von Markteinkommen: „all sources“
- Definition von Transfereinkommen: nicht vorhanden
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: Haushalt, Untergruppe „family“ (als Teil des Haushalts, jedoch nicht haushaltsübergreifend)
- offene Abfrage: nein
- kategoriale Abfrage: in 12 Kategorien, Schrittweite: 1000
- gemessen vs. geschätzt (durch wen?): Befragtenantwort
- monetär vs. nicht-monetär: es werden Einnahmen monetär in Landeswährung erhoben
- Erläuterung/Definition von Einkommen: keine
- Umfang der Abfrage/Anzahl der Fragen: 1 Frage

American „Panel Study of Income Dynamics“ - PSID:

- selbständig Erwerbstätige: „own a business“ oder Farmer/Rancher
- Bruttoeinkommen:
- für nicht-selbständig Erwerbstätige: „before taxes or other things“
- für selbständig Erwerbstätige: „share the income from the business“ (Entnahmen) und Profit
- zeitlicher Bezug: Jahr und Monat
- Definition von Markteinkommen: sehr differenziert nach diversen Einkommenstypen
- Definition von Transfereinkommen: sehr differenziert nach diversen Einkommenstypen
- Fokussierung auf Einkommenstyp: nein
- Individualeinkommen vs. Haushaltseinkommen: Individualeinkommen
- offene Abfrage: ja
- kategoriale Abfrage: nein
- gemessen vs. geschätzt (durch wen?): Befragtenantwort
- monetär vs. nicht-monetär: monetäre Einnahmen in Landeswährung

- Erläuterung/Definition von Einkommen: ausreichend durch separates Ansprechen jeden Typs
- Umfang der Abfrage/Anzahl der Fragen: pro Individuum 48 Fragen plus Unterfragen, durchzuführen für alle Individuen im Haushalt

WENN „MEHR“ NICHT GLEICHBEDEUTEND MIT „BESSER“ IST: AUSSCHÖPFUNGSQUOTEN UND STICHPROBENVERZERRUNGEN IN ALLGEMEINEN BEVÖLKERUNGSUMFRAGEN

ACHIM KOCH

Wenn Ausfälle bei der Realisierung von Zufallsstichproben nicht zufällig erfolgen, gefährden sie die Generalisierbarkeit von Umfrageergebnissen. In der Praxis wird häufig angenommen, daß eine Stichprobe um so weniger verzerrt ist, je höher die Ausschöpfung der Umfrage ausfällt. In dem Aufsatz wird auf Basis von sechs allgemeinen Bevölkerungsumfragen, deren Ausschöpfung zwischen 50 und 80 Prozent variiert, analysiert, inwiefern diese Annahme gerechtfertigt ist. Ein Vergleich der Randverteilungen ausgewählter soziodemographischer Merkmale mit Daten des Mikrozensus ergibt keine empirischen Belege dafür, daß die Stichprobenverzerrungen der Umfragen mit höherer Ausschöpfung geringer sind. Als eine mögliche Erklärung für diesen Befund wird diskutiert, inwiefern die von Umfrageinstituten berichteten Ausschöpfungsquoten zuverlässig sind.

Nonresponses in sample surveys jeopardize the generalizability of survey results, if they do not occur completely randomly. It is often assumed that the extent of nonresponse bias increases in line with nonresponse rates. The article discusses data from six general population surveys with response rates varying between 50 and 80 percent to see whether this, in fact, is the case. The marginals of several sociodemographic variables were compared with official microcensus data. No empirical evidence was found of a straightforward relationship between response rates and the extent of nonresponse bias. The reliability of survey organization reports of response rates is discussed as a possible explanation.

1. Einleitung

Bei Umfragen, die auf Zufallsstichproben basieren, gelingt es so gut wie nie, mit allen ausgewählten Zielpersonen ein Interview durchzuführen. Ein Teil der Personen wird während der Feldzeit nicht erreicht, ein Teil wird zwar kontaktiert, ist aber nicht fähig oder nicht bereit, an der Umfrage teilzunehmen. Wenn sich diese Ausfälle nicht zufällig über die Zielpopulation verteilen, liegt eine Stichprobenverzerrung bzw. ein Nonresponse-Bias vor, durch den die Generalisierbarkeit der Umfrageergebnisse beeinträchtigt wird.

In der Praxis wird häufig die Ausschöpfungsquote als Proxy für eine Messung des Nonresponse-Bias herangezogen. Man nimmt dabei an, daß die Verzerrung um so geringer ist, je höher die Teilnahmequote ausfällt.¹⁾ Dies ist aber nur begrenzt richtig, denn die Ausschöpfungsquote stellt lediglich eine Komponente für die Stärke des Nonresponse-Bias dar.

Bei linearen Statistiken wie Mittelwerten oder Anteilswerten ist der Nonresponse-Bias eine multiplikative Funktion der Höhe der Ausfallquote einerseits und der Differenz zwischen Teilnehmern und Nichtteilnehmern der Umfrage bei der betrachteten Variablen andererseits (vgl. Groves/Couper 1992: 103).

Der Nonresponse-Bias eines Anteilswertes z.B. berechnet sich nach

$$\text{Ausfallquote} * (p_T - p_{NT}),$$

wobei p_T dem Anteilswert der Teilnehmer und p_{NT} dem Anteilswert der Nichtteilnehmer bei der betreffenden Variablen entspricht. Die Formel verdeutlicht zweierlei:

1. Aus der Höhe der Ausfallquote kann nur auf das mögliche *Potential* eines Nonresponse-Bias geschlossen werden: je geringer die Ausschöpfung, desto größer kann die Verzerrung ausfallen. Prinzipiell ist es jedoch auch bei einer nur geringen Ausschöpfung möglich, daß der Bias klein ist, nämlich dann, wenn die Ausfälle weitgehend zufällig erfolgen, dh. wenn die Unterschiede zwischen Teilnehmern und Nichtteilnehmern bei einer Variablen gering sind.

2. Der Nonresponse-Bias ist eine *merkmalspezifische* Größe. Eine schlecht ausgeschöpfte Umfrage kann bei einem Merkmal eine deutliche Verzerrung aufweisen (weil sich Teilnehmer und Nichtteilnehmer hinsichtlich der betreffenden Variablen unterscheiden), bei einem anderen Merkmal jedoch weitgehend unverzerrt sein (weil Teilnehmer und Nichtteilnehmer bezüglich dieser Variablen nicht oder nur wenig differieren).

Will man der Frage des Nonresponse-Bias empirisch nachgehen, steht man vor dem Problem, daß die beiden Bestandteile dieser Verknüpfung in unterschiedlichem Maße einer Messung zugänglich sind: während die Ausfallquote - zumindest theoretisch - gut und eindeutig ermittelbar ist, liegen Informationen über die zweite Komponente - den Unterschied zwischen Teilnehmern und Nichtteilnehmern - normalerweise nicht vor. Die Abwesenheit von Informationen über die Nichtteilnehmer ist die Ursache dafür, daß eine routinemäßige Messung des Nonresponse-Bias kaum möglich ist.

Es gibt verschiedene Möglichkeiten, sich dem Problem der Ermittlung des Nonresponse-Bias zu nähern.²⁾ Eine häufig verfolgte Methode ist der Vergleich von Umfrageergebnissen mit Informationen der amtlichen Statistik. Bei Erhebungen wie der Volkszählung oder dem Mikrozensus spielt das Problem des Nonresponse aufgrund der Teilnahme-pflicht eine geringe Rolle. Im Mikrozensus etwa liegt der Unit-Nonresponse bei ca. drei Prozent (vgl. Lüttinger/Riede 1997: 27). Man kann deshalb annehmen, daß die Ergebnisse des Mikrozensus ein einigermaßen valides Außenkriterium zur Beurteilung möglicher Stichprobenverzerrungen von Umfragedaten darstellen.³⁾

Ich werde im folgenden die Randverteilungen ausgewählter Merkmale des Mikrozensus mit Daten aus sechs allgemeinen Bevölkerungsumfragen vergleichen, deren berichtete Ausschöpfung zwischen 50 und 80 Prozent beträgt. Aufgrund der begrenzten Zahl der im Mikrozensus verfügbaren Merkmale muß sich der Vergleich auf einige soziodemographische Variablen beschränken. Leitfragen der Analyse sind:

1. Welche Personengruppen sind in allgemeinen Bevölkerungsumfragen unter- bzw. überrepräsentiert?
2. Ist die Richtung der Verteilungsabweichungen in den betrachteten Umfragen ähnlich?
3. Gibt es empirische Anhaltspunkte dafür, daß der Nonresponse-Bias in Umfragen mit höherer Ausschöpfung niedriger ist als in solchen mit schlechterer Ausschöpfung?

2. Die analysierten Umfragen

Die Datenbasis bilden sechs allgemeine Bevölkerungsumfragen, die alle für sich in Anspruch nehmen, hohen Qualitätsstandards zu genügen. Fünf der betrachteten Umfragen sind sozialwissenschaftliche Surveys, die von ZUMA betreut wurden. Es sind dies der ALLBUS 1992 und 1994, der Wohlfahrtssurvey 1993 sowie zwei Umfragen des Sozialwissenschaften-Bus aus dem Jahre 1993. Berücksichtigt wird außerdem die Presse-medientranche der Media-Analyse 1994.⁴⁾ Alle sechs Surveys wurden in Form persön-

lich-mündlicher Interviews durchgeführt, die durchschnittliche Interviewdauer lag zwischen 50 (Media-Analyse) und 80 Minuten (ALLBUS 1992). In Übersicht 1 sind die wichtigsten Informationen zu Grundgesamtheit, Stichprobe und Ausschöpfung der Umfragen zusammengestellt.

Die Definition der Grundgesamtheit variiert etwas zwischen den einzelnen Umfragen. Alle sechs Surveys beziehen sich auf die Bevölkerung in Privathaushalten; Personen, die in Anstalten leben, wurden nicht befragt. Der Wohlfahrtssurvey und die beiden Umfragen des Sozialwissenschaften-Bus beschränken sich - wie die meisten sozialwissenschaftlichen Umfragen - auf die erwachsene deutsche Bevölkerung in Privathaushalten. In den beiden ALLBUS-Erhebungen dagegen wurden zusätzlich zur deutschen Bevölkerung auch die in Deutschland lebenden Ausländer befragt - dies allerdings nur dann, wenn das Interview in deutscher Sprache durchgeführt werden konnte. In der Media-Analyse wurden nicht nur erwachsene Deutsche, sondern auch Jugendliche zwischen 14 und 17 Jahren interviewt.

In fünf der sechs betrachteten Umfragen wurde das dreistufige Stichprobendesign des Arbeitskreises Deutscher Marktforschungsinstitute („ADM-Stichprobensystem“) eingesetzt, das in der Bundesrepublik bei der Ziehung von Zufallsstichproben für persönlich-mündliche Bevölkerungsumfragen eine dominierende Stellung einnimmt.⁵⁾ Bei diesem Verfahren wird in der ersten Stufe eine Stichprobe von Wahlbezirken gezogen. In der zweiten Auswahlstufe ermitteln die Interviewer in den ausgewählten Wahlbezirken anhand eines zufällig ausgewählten Startpunktes und vorgegebener Begehungsregeln eine festgelegte Zahl von Haushalten.⁶⁾ In der dritten Ziehungsstufe wählen die Interviewer mit Hilfe eines vorbereiteten Zufallsschemas („Schwedenschlüssel“) die zu befragende Person aus den zur Grundgesamtheit der Umfrage zählenden Haushaltsmitgliedern aus.⁷⁾

Eine Alternative zum ADM-Design, die in der Praxis allerdings nur selten gewählt wird, stellt die im ALLBUS 1994 erfolgte Stichprobenziehung aus Einwohnermelderegistern dar. Bei diesem zweistufigen Verfahren wird zunächst eine Stichprobe von Gemeinden gebildet; im ALLBUS 1994 waren es 151 Gemeinden. In der zweiten Auswahlstufe wurden anschließend aus den Einwohnermelderegistern dieser Gemeinden jeweils 40 Personen für die Befragung zufällig ausgewählt. Die Interviewer erhielten die zu befragenden Personen mit Namen und Adresse vorgegeben. Die Phase der Stichprobenbildung ist bei diesem Verfahren damit vollständig von der Feldarbeit getrennt.

Übersicht 1: Die analysierten Umfragen

	ALLBUS 1992	ALLBUS 1994	Wohlfahrts- survey 1993	Sozialwissen- schaften-Bus III 1993*	Sozialwissen- schaften-Bus II 1993*	Media-Analyse (Pressemedientranche) 1994
Umfrageinstitut	Infratest			GFM-GETAS		Basis Research, GFM-GETAS, Ifak, Marplan, Media Markt Analysen, Sample
Grundgesamtheit	Erwachsene Bevölkerung in Privathaushalten		Erwachsene deutsche Bevölkerung in Privathaushalten			Deutsche Bevölkerung in Privathaushalten im Alter von 14 und mehr Jahren
Stichprobenverfahren	ADM-Stichprobe „Standard-Random“	Einwohnermelde- register- Stichprobe	ADM-Stichprobe „Random-Route“			ADM-Stichprobe „Adress-Random“
Netto-Fallzahl						
West	2400	2342	2046	2024	2143	19247
Ost	1148	1108	1016	2142	1092	6374
Ausschöpfung (in %)						
West	51,9	53,2	62,9	68,0	73,0	80,2
Ost	54,7	55,2	64,4	71,1	72,9	80,2

* Der Sozialwissenschaften-Bus I 1993 wurde nicht berücksichtigt, weil er lediglich in Ostdeutschland durchgeführt wurde.

Sowohl das ADM-Design als auch der Zugang über die Einwohnermelderegister führen zu Zufallsstichproben der Bevölkerung: der Unterschied besteht allerdings darin, daß bei der Einwohnermelderegister-Stichprobe die Befragungspersonen eine gleiche Auswahlchance haben, während dies beim ADM-Design nur für die Befragungshaushalte gilt. Die Auswahlchance der einzelnen Haushaltsmitglieder ist hier um so geringer, je größer die Zahl der Haushaltspersonen ist, die zur Grundgesamtheit der betreffenden Umfrage zählen. Dies hat zur Folge, daß bei Auswertungen auf Personenebene vom Design her eine Gewichtung mit der - um die nicht zur Grundgesamtheit zählenden Haushaltsmitglieder reduzierten - Haushaltsgröße vorgesehen ist (exakter: mit einem Faktor, der dem reziproken Wert der Auswahlchance entspricht).

In den fünf sozialwissenschaftlichen Umfragen wurden jeweils ungefähr 2.000 bis 2.400 Personen in Westdeutschland und 1.000 Personen in Ostdeutschland befragt. Lediglich im Sozialwissenschaften-Bus III 1993 wurden auch im Osten ca. 2.000 Fälle realisiert. Die Media-Analyse dagegen weist deutlich höhere Fallzahlen auf: in dieser Gemeinschaftsuntersuchung, an deren Durchführung insgesamt sechs Umfrageinstitute beteiligt waren, wurden rund 19.000 Fälle im Westen und 6.000 im Osten realisiert. Jedes beteiligte Umfrageinstitut hat ca. 3.000 Interviews in Westdeutschland und 1.000 in Ostdeutschland durchgeführt.

In den Ausschöpfungsquoten der Studien, wie sie von den Umfrageinstituten berichtet werden, bestehen große Unterschiede.⁸⁾ Die beiden ALLBUS-Umfragen weisen mit etwas über 50 Prozent die niedrigste Ausschöpfung auf, es folgt der Wohlfahrtssurvey mit etwas über 60 Prozent, dann die beiden Umfragen des Sozialwissenschaften-Bus mit ungefähr 70 Prozent. Spitzenreiter ist die Media-Analyse mit 80 Prozent. West- und Ostdeutschland unterscheiden sich bei den einzelnen Umfragen kaum in der Ausschöpfung.⁹⁾

3. Vergleich von Umfrageergebnissen mit Mikrozensusdaten

Um zu untersuchen, ob diese unterschiedlichen Ausschöpfungsquoten einen Einfluß auf den Nonresponse-Bias der Umfragen haben, werden im folgenden die Randverteilungen ausgewählter demographischer Merkmale mit Ergebnissen des Mikrozensus verglichen. Der Mikrozensus ist eine seit 1957 (nahezu) jedes Jahr stattfindende amtliche Erhebung, bei der die Erfassung erwerbsstatistischer und soziodemographischer Merkmale im Mittelpunkt steht. Der Erhebung liegt eine Ein-Prozent-Haushaltsstichprobe der wohnberechtigten Bevölkerung zugrunde. Sie umfaßt seit 1991 etwa 730.000 befragte Personen in 330.000 Haushalten im wiedervereinigten Deutschland. Für die ausgewählten Personen und Haushalte besteht Teilnahmepflicht (vgl. Lüttinger/Riede 1997).

Die vorliegende Analyse basiert auf der ZUMA zur Verfügung stehenden 70-Prozent-Unterstichprobe des Mikrozensus 1993. Mit der Wahl dieses Mikrozensus nehme ich in Kauf, daß die Erhebungszeitpunkte der beiden ALLBUS-Umfragen jeweils um ein Jahr abweichen.¹⁰⁾ Da sich die Verteilungen sozialstruktureller Merkmale normalerweise nur langsam ändern, erscheint dies vertretbar.

In den Vergleich einbezogen wurden soziodemographische Merkmale, für die bereits in früheren Untersuchungen Verteilungsabweichungen zwischen Mikrozensus und Umfragedaten festgestellt wurden. Hartmann/Schimpl-Neimanns (1992) fanden bei einem Vergleich der ALLBUS-Erhebungen 1986, 1988 und 1990 mit Daten des Mikrozensus, daß in den Umfragedaten Frauen, über 70jährige, Personen mit Hauptschulabschluß, Erwerbstätige, Arbeiter, Verwitwete, Geschiedene und Ledige sowie Alleinlebende unterrepräsentiert sind (vgl. Übersicht 2).¹¹⁾ In der Übersicht kursiv dargestellt sind dabei Ergebnisse, die als relativ gut belegt gelten können, weil sie auch in anderen Untersuchungen zur „Demographie“ von Ausfällen wiederholt beobachtet wurden (vgl. z.B. die Zusammenfassung empirischer Befunde bei Burton 1996: 13ff., Esser u.a. 1989: 111ff., Goyder 1987: 80ff., Groves 1989: 201ff.). Dies betrifft die Untererfassung von Älteren, Personen mit niedrigem Bildungsabschluß sowie Arbeitern (These des Mittelschichtbias) und alleinlebenden Personen.

Übersicht 2: Untersuchte Merkmale

Merkmal	Unterrepräsentiert sind
Geschlecht	Frauen
Alter	<i>über 70jährige</i>
Schulabschluß	<i>Personen mit Hauptschulabschluß</i>
Stellung im Erwerbsleben	Erwerbstätige
Berufliche Stellung	<i>Arbeiter</i>
Familienstand	Verwitwete, Geschiedene, Ledige
Haushaltsgröße	<i>Alleinlebende</i>

Damit Verteilungsabweichungen zwischen amtlicher Statistik und Umfragedaten als Ausdruck eines Nonresponse-Bias interpretiert werden dürfen, müssen die zugrundeliegenden Zielpopulationen und Erhebungskonzepte vergleichbar sein. Um die Bevölkerungskonzepte zu vereinheitlichen, basieren die folgenden Auswertungen auf der erwachsenen deutschen Bevölkerung in Privathaushalten. Zur Erzielung einer weitgehend äquivalenten Operationalisierung der Merkmale wurden die Variablen jeweils auf einen gemeinsamen Standard rekodiert. Die vorgenommenen Rekodierungen für die Merkmale Schulabschluß, Stellung im Erwerbsleben und berufliche Stellung, die

etwas komplexeren Charakter haben, sind im Anhang dokumentiert. Um zu vermeiden, daß zufällige Differenzen aufgrund des Stichprobenfehlers fälschlicherweise als Folge eines Nonresponse-Fehlers interpretiert werden, werde ich im folgenden nur Abweichungen, die bei allen (oder nahezu allen) Umfragen *in gleicher Richtung* auftreten, als Beleg für eine Stichprobenverzerrung werten.¹²⁾ Eine zusätzliche Absicherung erfährt diese Interpretation immer dann, wenn die Abweichungen mit dem gegebenen Vorwissen über die Systematik von Ausfällen übereinstimmen.

4. Ausschöpfungsquoten und Stichprobenverzerrungen

Die folgenden Tabellen enthalten jeweils die Merkmalsverteilung für den Mikrozensus. Für die einzelnen Surveys werden die Prozentsatzdifferenzen zum Mikrozensus wiedergegeben. Negative Abweichungen sind zur besseren Orientierung kursiv dargestellt. Die Umfragen sind nach der Ausschöpfung in aufsteigender Reihenfolge angeordnet.

Sämtliche Tabellen basieren auf der erwachsenen deutschen Bevölkerung in Privathaushalten. Für die Studien nach dem ADM-Verfahren wurden die designgewichteten Daten verwendet.¹³⁾

Geschlecht

In Westdeutschland stimmt die Verteilung nach dem Merkmal Geschlecht in fast allen Umfragen mit den Daten des Mikrozensus recht gut überein. Eine Ausnahme bildet lediglich der ALLBUS 1994, bei dem der Anteil der Frauen um 3,8 Prozentpunkte unter dem Wert des Mikrozensus liegt.¹⁴⁾

Tabelle 1: Differenz zum Mikrozensus beim Merkmal Geschlecht (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Männlich	47,2	1,8	3,8	0,2	-1,0	1,3	-0,3
Weiblich	52,8	-1,8	-3,8	-0,2	1,0	-1,3	0,3
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Männlich	46,7	2,2	1,4	4,7	2,4	3,9	1,5
Weiblich	53,3	-2,2	-1,4	-4,7	-2,4	-3,9	-1,5

In Ostdeutschland sind beim Geschlecht die Differenzen zum Mikrozensus etwas größer als in Westdeutschland. In allen sechs Umfragen liegt der Anteil der Frauen unter dem Ergebnis des Mikrozensus, am stärksten ist die Abweichung im Wohlfahrtssurvey mit -4,7 Prozentpunkten, am geringsten beim ALLBUS 1994 mit -1,4 Prozentpunkten. Die Gleichförmigkeit der Abweichung über alle sechs Umfragen stellt ein Indiz für einen - schwach ausgeprägten - Nonresponse-Bias dar.¹⁵⁾ Hinweise darauf, daß die Verzerrung mit steigender Ausschöpfung geringer wird, finden sich nicht.

Alter

In Westdeutschland ist die Unterrepräsentierung der über 70jährigen das hervorstechende Ergebnis bei der Altersverteilung. In allen sechs Umfragen ist der Anteil dieser Altersgruppe um 3,2 bis 4,5 Prozentpunkte niedriger als im Mikrozensus. Im Gegenzug sind in allen Umfragen die 30- bis 49jährigen stärker als im Mikrozensus vertreten. Der überproportionale Ausfall älterer Personen steht in Einklang mit den bekannten Regelmäßigkeiten des Teilnahmeverhaltens, man kann hier also mit großer Wahrscheinlichkeit annehmen, daß ein Nonresponse-Effekt vorliegt. Eine bessere Anpassung an den Mikrozensus mit zunehmender Ausschöpfung der Umfragen ist nicht festzustellen.

Tabelle 2: Differenz zum Mikrozensus beim Merkmal Alter (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
18 bis 29 Jahre	20,8	0,4	0,2	-2,6	-0,5	1,7	-1,6
30 bis 39 Jahre	18,1	3,0	2,7	2,0	2,9	1,1	1,6
40 bis 49 Jahre	15,3	1,9	0,4	1,1	1,3	2,4	2,7
50 bis 59 Jahre	18,3	-0,2	0,5	1,8	-1,0	-1,0	0,3
60 bis 69 Jahre	13,8	-0,5	-0,3	1,4	0,8	0,4	0,2
70 Jahre und älter	13,7	-4,5	-3,6	-3,7	-3,5	-4,5	-3,2
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
18 bis 29 Jahre	19,7	-1,7	-1,7	-7,2	-5,2	0,1	-4,0
30 bis 39 Jahre	19,9	2,8	-1,5	0,8	1,0	-1,1	1,1
40 bis 49 Jahre	15,9	2,2	2,2	2,3	2,2	3,2	2,7

50 bis 59 Jahre	19,4	3,5	1,6	3,8	2,3	0,8	1,8
60 bis 69 Jahre	13,3	-1,4	0,7	4,1	3,4	1,8	3,1
70 Jahre und älter	11,8	-5,4	-1,3	-3,8	-3,8	-4,9	-4,7

Auch in Ostdeutschland ist in allen Umfragen eine Unterrepräsentierung der über 70jährigen zu beobachten. Diese fällt zumeist in ähnlicher Größenordnung aus wie in Westdeutschland.¹⁶⁾ In drei Umfragen ist darüber hinaus die jüngste Altersgruppe der 18- bis 29jährigen deutlich zu gering vertreten: dies betrifft die Media-Analyse (-4,0 Prozentpunkte), den Sozialwissenschaften-Bus III 1993 (-5,2 Prozentpunkte) und den Wohlfahrtssurvey 1993 (-7,2 Prozentpunkte). Die Stärke der Abweichung und die z.T. auch in anderen Studien berichtete höhere Ausfallquote jüngerer Altersgruppen (vgl. z.B. Burton 1996: 15) legen es nahe, auch diese Abweichung als systematische Stichprobenverzerrung zu werten. Im Vergleich zum Mikrozensus überrepräsentiert sind im Osten in (nahezu) allen Umfragen die 40- bis 69jährigen. Insgesamt betrachtet weist der ALLBUS 1994 mit lediglich 55 Prozent Ausschöpfung die beste Übereinstimmung mit dem Mikrozensus auf, bei allen anderen Umfragen sind die Abweichungen deutlich größer.¹⁷⁾

Schulabschluss

Beim Merkmal Schulabschluss ist nach der Literatur eine Unterrepräsentierung von Personen mit niedrigem Bildungsabschluss zu erwarten. In Westdeutschland trifft dies für alle sechs Umfragen zu: der Anteil der Personen mit Volks-/Hauptschulabschluss liegt konsistent unter dem Wert des Mikrozensus.¹⁸⁾ Auf der anderen Seite sind Personen mit mittleren Bildungsabschlüssen und - in etwas schwächerem Maße - auch Personen mit Abitur oder Hochschulbildung (Ausnahme: Media-Analyse) in den Umfragen überrepräsentiert. Den größten Unterschied zum Mikrozensus weist die Umfrage mit der zweithöchsten Ausschöpfung auf: im Sozialwissenschaften-Bus II 1993 liegt der Anteil der Volks- und Hauptschulabsolventen um 10,3 Prozentpunkte unter dem Wert des Mikrozensus. Am besten stimmt die Media-Analyse mit den Daten der amtlichen Statistik überein.

In Ostdeutschland kann aufgrund der Unterschiede in den Abfragetechniken lediglich zwischen Abitur/Fachhochschule/Hochschule auf der einen Seite und allen niedrigeren Abschlüssen auf der anderen Seite differenziert werden (vgl. Anhang).¹⁹⁾ Bei dieser Differenzierung zeigt sich eine systematische Überrepräsentierung von Personen mit Abitur bzw. Fachhoch-/Hochschulabschluss, während Personen mit niedrigeren Abschlüssen in allen Umfragen zu gering vertreten sind. Die Differenzen zum Mikrozensus sind mit 4,3 bis 6,3 Prozentpunkten annähernd gleich groß, eine Abnahme der Verzerrung mit zunehmender Ausschöpfung der Umfragen ist nicht festzustellen.

Tabelle 3: Differenz zum Mikrozensus beim Merkmal Schulabschluß (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Bis Volks-/Hauptschul- abschluß	59,6	-6,2	-5,7	-7,6	-6,0	-10,3	-3,5
Mittlere Reife, Fachhochschulreife	23,5	3,7	4,2	5,1	5,1	7,3	3,8
Abitur, Fachhochschule, Hochschule	16,9	2,5	1,5	2,5	0,9	3,1	-0,2
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Bis Volks-/Hauptschul- abschluß, Mittlere Reife, Fachhochschulreife	86,0	-6,1	-4,3	-5,0	-6,2	-4,3	-5,0
Abitur, Fachhochschule, Hochschule	13,9	6,2	4,3	5,1	6,3	4,4	5,1

Tabelle 4: Differenz zum Mikrozensus beim Merkmal Stellung im Erwerbsleben (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Hauptberuflich erwerbstätig	52,2	-0,9	4,4	-1,7	-3,9	-1,8	2,8
Nicht hauptberuflich erwerbstätig	47,8	0,9	-4,4	1,7	3,9	1,8	-2,8
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Hauptberuflich erwerbstätig	52,9	4,9	0,2	-2,7	-6,5	-4,3	-1,5
Nicht hauptberuflich	47,1	-4,9	-0,2	2,7	6,5	4,3	1,5

erwerbstätig

Stellung im Erwerbsleben

Die Stellung im Erwerbsleben ist unter den analysierten Merkmalen dasjenige, bei dem Unterschiede in den Erhebungskonzepten am ehesten die Vergleichbarkeit beeinträchtigen (vgl. die Dokumentation im Anhang). In den gewählten Abgrenzungen sind die Anteile der Erwerbstätigen in den sechs Umfragen in West und Ost z.T. größer, z.T. kleiner als im Mikrozensus. Vor dem Hintergrund unterschiedlicher Erhebungskonzepte kann hier keine Aussage zum Nonresponse-Bias gemacht werden.

Berufliche Stellung

Was die berufliche Stellung betrifft, ist nach der Literatur mit einer Untererfassung der Arbeiter zu rechnen. Diese läßt sich in Westdeutschland auch bei allen Umfragen beobachten, allerdings sind die Abweichungen zumeist gering. Im Gegenzug sind in allen Surveys die Angestellten zu stark vertreten - am ausgeprägtesten im Sozialwissenschaftler-Bus II 1993 und in der Media-Analyse. Die beiden am besten ausgeschöpften Umfragen weisen damit auch insgesamt gesehen die größte Differenz zum Mikrozensus auf, während man den anderen Surveys bei diesem Merkmal eine relativ gute Übereinstimmung mit den amtlichen Daten attestieren kann.

Tabelle 5: Differenz zum Mikrozensus beim Merkmal Berufliche Stellung (in Prozentpunkten)*

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Arbeiter	30,2	-2,0	-0,5	-2,7	-0,7	-4,1	-2,5
Angestellte	46,5	1,7	0,9	2,3	2,3	5,4	6,1
Beamte	8,8	1,9	-0,5	1,2	-0,8	-0,1	-2,2
Selbständige	11,0	-1,0	1,6	-0,5	-0,5	-0,7	0,1
In Ausbildung	3,5	-0,7	-1,4	-0,3	-0,3	-0,4	-1,4
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Arbeiter	39,1	-3,7	-0,4	-2,6	-4,3	-6,8	-2,4
Angestellte	48,6	3,4	-4,0	0,9	-0,1	-0,1	-1,9
Beamte	1,8	-0,6	1,1	-0,4	1,0	0,8	0,9
Selbständige	6,7	1,7	4,0	2,7	3,8	5,8	3,8

In Ausbildung	3,8	-0,8	-0,7	-0,6	-0,4	0,3	-0,4
---------------	-----	------	------	------	------	-----	------

* Derzeit hauptberuflich erwerbstätige Personen.

Auch in Ostdeutschland sind Arbeiter unterrepräsentiert, am stärksten in den beiden Erhebungen des Sozialwissenschaften-Bus (-4,3 bzw. -6,8 Prozentpunkte). Dafür ist in allen Umfragen der Anteil der Selbständigen besonders hoch. In Relation zu ihrem Anteil an den Erwerbstätigen ist die Abweichung vom Mikrozensus sehr groß: im Sozialwissenschaften-Bus II 1993 z.B. ist der Anteil der Selbständigen beinahe doppelt so hoch wie im Mikrozensus.²⁰⁾ Insgesamt läßt sich auch in Ostdeutschland keine bessere Anpassung an den Mikrozensus mit einer Zunahme der Ausschöpfung feststellen.

Familienstand

In Einklang mit bisherigen Befunden läßt sich beim Merkmal Familienstand eine Unterschätzung nicht-verheirateter Personen beobachten. In West- und Ostdeutschland liegt der Anteil der Verwitweten/Geschiedenen sowie der Ledigen in allen Umfragen jeweils unter dem Wert des Mikrozensus, der Anteil der Verheirateten darüber.²¹⁾

Tabelle 6: Differenz zum Mikrozensus beim Merkmal Familienstand (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Verheiratet	60,5	7,3	2,8	10,1	6,3	5,0	7,4
Verwitwet/Geschieden	15,2	-3,7	-1,1	-4,4	-3,7	-4,6	-2,6
Ledig	24,3	-3,6	-1,8	-5,7	-2,6	-0,4	-4,8
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Verheiratet	64,3	10,5	3,7	12,2	8,9	2,9	8,9
Verwitwet/Geschieden	16,9	-6,2	-0,5	-4,8	-3,7	-3,0	-5,1
Ledig	18,8	-4,2	-3,3	-7,4	-5,1	0,1	-3,9

Eine geringe Differenz zum Mikrozensus besteht in beiden Teilen Deutschlands beim ALLBUS 1994, eine sehr große Abweichung beim Wohlfahrtssurvey. In dieser Umfrage liegt z.B. der Anteil der Verheirateten jeweils um mehr als zehn Prozentpunkte über dem Ergebnis des Mikrozensus. Weder in West- noch in Ostdeutschland werden die Unterschiede zum Mikrozensus mit einer Zunahme der Ausschöpfung geringer. So weist auch die Media-Analyse trotz 80 Prozent Ausschöpfung beim Familienstand starke Verzerrungen auf.

Haushaltsgröße

Auch hinsichtlich des Merkmals Haushaltsgröße ergibt sich in der Grundtendenz ein einheitliches Bild. In West- und Ostdeutschland sind wie erwartet alleinlebende Personen in allen Umfragen unterrepräsentiert. Nur geringe Unterschiede zum Mikrozensus weist wiederum der ALLBUS 1994 auf,²²⁾ die größten Abweichungen bestehen beim ALLBUS 1992 und Wohlfahrtssurvey 1993.²³⁾ Auch die Verteilungen bei diesem Merkmal zeigen, daß aus der Höhe der Ausschöpfung keine direkten Rückschlüsse auf die Verzerrung der Stichprobe getroffen werden können.

Tabelle 7: Differenz zum Mikrozensus beim Merkmal Haushaltsgröße (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
1 Person	19,9	-7,3	-1,8	-8,4	-4,3	-5,0	-4,6
2 Personen	34,7	-1,2	0,3	-1,4	2,7	1,7	2,6
3 Personen	21,3	3,6	-0,6	2,8	0,1	1,2	0,6
4 Personen	16,8	2,7	0,9	5,0	0,7	2,3	1,3
5 und mehr Personen	7,3	2,1	1,1	1,9	0,8	-0,2	0,1
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
1 Person	15,9	-7,8	-2,2	-7,1	-5,9	-3,1	-5,3
2 Personen	34,6	0,7	-0,3	1,2	4,7	3,6	2,8
3 Personen	24,3	1,4	0,9	4,2	0,8	-0,8	1,7
4 Personen	20,0	3,3	1,1	1,3	-0,5	-0,1	-0,2
5 und mehr Personen	5,1	2,5	0,7	0,5	1,1	0,5	1,0

5. Fazit und Schlußbemerkungen

Als Ergebnis der durchgeführten Analysen läßt sich festhalten:

- Die analysierten Umfragen weisen bei den betrachteten soziodemographischen Merkmalen in der Regel Abweichungen in gleicher Richtung vom Mikrozensus auf.
- Die Struktur der Abweichungen stimmt im großen und ganzen mit dem Wissen über die Systematik von Ausfällen überein, so daß es naheliegt, sie als Ausdruck eines Nonresponse-Bias zu interpretieren.

- Es finden sich keine empirischen Indizien dafür, daß der Bias mit der Höhe der Ausschöpfungsquote abnimmt. Zumindest hinsichtlich der betrachteten demographischen Merkmale gibt es keine Hinweise dafür, daß besser ausgeschöpfte Umfragen geringere Stichprobenverzerrungen aufweisen als Umfragen mit niedrigerer Ausschöpfung. Die Höhe der berichteten Ausschöpfung ist somit kein guter Indikator für die Größe des Nonresponse-Bias.²⁴⁾

Wie ist dieser letzte Befund erklärbar? Hierfür kommen im Prinzip zwei Möglichkeiten in Betracht. Eine mögliche Erklärung könnte darin liegen, daß in den besser ausgeschöpften Umfragen zwar die Ausschöpfung höher, gleichzeitig aber auch die Unterschiede zwischen Teilnehmern und Nichtteilnehmern größer sind als in den schlechter ausgeschöpften Umfragen (vgl. Groves 1989: 145ff.). Ein solcher Fall tritt dann ein, wenn im Zuge der Bemühungen zur Verbesserung der Ausschöpfung lediglich solche Personen doch noch für ein Interview gewonnen werden können, die eher „untypisch“ für die gesamte Gruppe der ursprünglichen Ausfälle sind. Feldarbeitsstrategien, die lediglich darauf abzielen, die Ausschöpfungsquote insgesamt zu maximieren, bieten keine Gewähr für eine Reduzierung des Nonresponse-Bias.²⁵⁾

Plausibler scheint mir jedoch ein anderes Erklärungsmodell zu sein. Zu fragen ist nämlich, inwieweit die berichteten Ausschöpfungsergebnisse - insbesondere die hohen Werte von 70 Prozent und mehr - die Realität korrekt wiedergeben. Womöglich resultiert der fehlende Zusammenhang zwischen der Höhe der Ausschöpfung und dem Ausmaß des Nonresponse-Bias daraus, daß die betrachteten Umfragen sich de facto in der Ausschöpfung doch nicht so sehr unterscheiden, wie die berichteten Quoten es nahelegen.

Bei den analysierten Umfragen kann lediglich im ALLBUS 1994 die Ausschöpfung (weitgehend) zuverlässig berechnet werden, weil aufgrund der Ziehung einer Stichprobe von Personenadressen aus Einwohnermelderegistern vor Feldbeginn ein genau definiertes Stichprobenbrutto vorlag (zu den durchgeführten Kontrollen vgl. Koch 1995). Sämtliche anderen Umfragen dagegen wurden nach dem ADM-Design durchgeführt. Die faktische Bruttostichprobe wird hier durch die Interviewer in der zweiten und dritten Auswahlstufe gebildet. Die Zuverlässigkeit der Ausschöpfungsberechnung hängt davon ab, wie exakt die Interviewer die Vorgaben beachten. Bei diesem Verfahren ist nicht auszuschließen, daß Interviewer die strikte Einhaltung der Random-Anweisungen zugunsten der Realisierung einer möglichst großen Zahl von Interviews opfern und *undokumentiert* schwer erreichbare oder kooperationsunwillige Zielpersonen gegen leichter erreichbare und kooperative Personen *substituieren*.²⁶⁾

Ein einfaches Rechenexempel verdeutlicht, daß bereits relativ „kleine“ Manipulationen die Höhe der berichteten Ausschöpfung merklich beeinflussen können: Angenommen, ein Interviewer hätte in einem Wahlbezirk acht Adressen als Brutto vorgegeben und realisierte daraus unter Beachtung aller Regeln fünf Interviews, dann läge seine Ausschöpfung bei 62,5 Prozent. Erzielte er regelwidrig noch ein weiteres Interview mit einer Person, die nicht zu seinem Brutto gehört, könnte er seine Ausschöpfung bereits auf 75 Prozent „verbessern“!

Inwieweit ein solches Fehlverhalten bei den vorliegenden Umfragen - und in der Umfragepraxis generell - eine Rolle spielt, kann hier nicht beurteilt werden. Eine öffentliche Diskussion darüber, welche Ausschöpfungen heutzutage faktisch möglich sind, findet in der Bundesrepublik nicht oder nur in Ansätzen statt (vgl. z.B. Porst 1996). Nur vereinzelt finden sich kritische Äußerungen wie die von Sommer (1987), der generell bezweifelt, daß in allgemeinen Bevölkerungsumfragen Ausschöpfungsquoten von 70 Prozent und mehr im normalerweise gegebenen Zeit- und Kostenrahmen überhaupt noch realisierbar sind.²⁷⁾

Eine wichtige Voraussetzung für realistische Ausschöpfungsberechnungen ist, daß Umfrageinstitute und Auftraggeber Abschied nehmen von der Vorstellung, daß die Ausschöpfungsquote alleine einen Indikator für die Stichprobenqualität darstellt und die Erzielung eines bestimmten Ausschöpfungslevels quasi automatisch „Qualität“ garantiert. Es wäre wünschenswert, wenn die vorliegenden Ergebnisse dazu beitragen würden, diese Annahme zu relativieren und den „Mythos der Ausschöpfung“ (Sommer 1987) durch eine breiter angelegte Diskussion über Qualitätskriterien für Stichproben zu ersetzen.

Anmerkungen

- 1) So heißt es etwa im Rahmenschema für Werbeträgeranalysen 1988 des Zentralverbands der deutschen Werbewirtschaft auf S. 15: „Je höher die Ausschöpfung, um so besser ist die Qualität der Stichprobe“ (zit. nach Landgrebe 1992: 20).
- 2) Einen Überblick über die Verfahren zur Gewinnung von Informationen über Ausfälle geben Groves (1989: 186ff.) und Schnell (1997: 134ff.).
- 3) Dies gilt natürlich nur dann, wenn nicht andere Fehlerquellen wie etwa Meßfehler die Gültigkeit der Mikrozensus-Daten beeinträchtigen. Hartmann (1990) fand eine gute Übereinstimmung der Randverteilungen demographischer Merkmale aus dem Mikrozensus mit anderen statistischen Quellen (wie z.B. der Volkszählung oder der Hochschulstatistik).

4) Nähere Informationen zu den einzelnen Umfragen bieten: ALLBUS 1992 (Braun u.a. 1993), ALLBUS 1994 (Koch u.a. 1994), Wohlfahrtssurvey 1993 (Infratest Sozialforschung 1993), Sozialwissenschaften-Bus (GFM-GETAS o.J.), Media-Analyse (Arbeitsgemeinschaft Media-Analyse o.J.).

5) Informationen zum ADM-Stichprobenverfahren finden sich in mehreren Beiträgen in Gabler/Hoffmeyer-Zlotnik (1997). Auch die im ALLBUS 1994 erfolgte Stichprobenziehung aus Einwohnermelderegistern ist dort ausführlich beschrieben.

6) In den einzelnen Umfragen kamen dabei unterschiedliche Verfahrensweisen zum Einsatz. In der Media-Analyse wurde das sog. Adress-Random-Verfahren verwendet, bei dem eine bestimmte Zahl von Haushaltsadressen in einem getrennten Arbeitsschritt vor der Durchführung der Interviews ermittelt wird. In jedem Wahlbezirk wurden durch die Interviewer zunächst 50 Haushaltsadressen aufgenommen und im Umfrageinstitut wurden fünf Adressen für die eigentliche Befragung zufällig ausgewählt. Im Wohlfahrtssurvey und in den beiden Umfragen des Sozialwissenschaften-Bus dagegen kam das in der Praxis gebräuchlichere Random-Route-Verfahren zum Einsatz, bei dem die Haushaltsauflistung in einem Zuge mit der Realisierung der Interviews erfolgt. Die Interviewer hatten entlang des Begehungsweges in systematischer Auswahl jeweils acht bzw. neun Haushalte zu kontaktieren und in möglichst jedem dieser Haushalte ein Interview durchzuführen. Auch das - im Sprachgebrauch von Infratest - als Standard-Random bezeichnete Vorgehen beim ALLBUS 1992 entspricht dem Random-Route-Modell. Der Unterschied zum herkömmlichen Random-Route-Verfahren besteht allerdings darin, daß den Interviewern nicht mehr ein festes Brutto von zu bearbeitenden Haushalten vorgeschrieben wird, sondern lediglich eine Nettofallzahl zu realisierender Interviews, die aus einem relativ großzügig bemessenen Brutto erzielt werden kann. Im ALLBUS 1992 sollten pro Wahlbezirk sechs (West) bzw. fünf (Ost) Interviews aus maximal 15 bearbeitbaren Adressen realisiert werden.

7) In allen fünf Surveys wurde jeweils eine Person im Haushalt befragt. Eine Ausnahme bildet die Media-Analyse, bei der in großen Haushalten (mit vier und mehr Personen im Alter von 14 und mehr Jahren) zwei Befragungspersonen ausgewählt wurden.

8) Die Berechnung der Ausschöpfungsquote erfolgte in einheitlicher Weise. Bei allen sechs Umfragen wurden Ausfälle, die nicht zur Grundgesamtheit der Umfrage zählten, sowie nicht auffindbare bzw. nicht bearbeitete Adressen nicht zum einsetzbaren Brutto gezählt (sog. „stichprobenneutrale Ausfälle“). Als ausschöpfungsrelevante, nicht-stichprobenneutrale Nonresponses wurden Ausfälle aufgrund von Nichterreichbarkeit, Krankheit, Verweigerung oder mangelnder Deutschkenntnisse gewertet.

9) Es ist nicht Ziel des vorliegenden Beitrages, die Ursachen für diese unterschiedlichen Ausschöpfungsergebnisse zu analysieren. Einfache Erklärungen dafür sind kaum möglich. So ist etwa auf Basis der verfügbaren Informationen kein Zusammenhang zwischen der Ausschöpfung und der Länge der Feldzeit, der

durchschnittlichen Interviewdauer oder den Studienkosten festzustellen. Ich werde im letzten Abschnitt diese Thematik nochmals aufgreifen.

10) Man beachte, daß bei der Media-Analyse 1994 der überwiegende Teil der Interviews im Jahr 1993 durchgeführt wurde. Um die Erhebungsjahre auch für die beiden ALLBUS-Umfragen konstant zu halten, hätte man natürlich auch die Mikrozensus 1992 und 1994 heranziehen können. Ich habe dies nicht getan, weil in beiden Erhebungen Bildungsmerkmale nicht erfaßt wurden.

11) Die Ergebnisse beruhen auf den - mit dem Design-Gewicht - personenrepräsentativ gewichteten Verteilungen.

12) Unter der Annahme einfacher Zufallsstichproben wären in Westdeutschland bei den fünf sozialwissenschaftlichen Umfragen folgende Differenzen zum Mikrozensus auf dem 5-Prozent-Niveau signifikant: Bei Anteilswerten von ungefähr 10 bzw. 90 Prozent Differenzen größer als ca. 1,5 Prozentpunkte, bei Anteilswerten nahe 50 Prozent Differenzen größer als ca. 2,5 Prozentpunkte. Bei der Media-Analyse liegen die Grenzen aufgrund der höheren Fallzahl niedriger: Hier wären bereits Abweichungen von einem halben bzw. einem Prozentpunkt signifikant. In Ostdeutschland lägen die Werte aufgrund der niedrigeren Fallzahlen mit 2,0 und 3,5 Prozentpunkten bei den sozialwissenschaftlichen Umfragen und 1,0 bzw. 1,5 Prozentpunkten bei der Media-Analyse etwas höher. Diese Zahlen sind als untere Grenze zu betrachten, da die Konfidenzintervalle von Schätzern auf der Basis komplexer Stichproben, wie sie hier vorliegen, zumeist größer sind als bei einfachen Zufallsstichproben (vgl. Schnell 1997: 65ff., 272ff.).

13) Wie oben erwähnt, werden beim ADM-Design lediglich die Haushalte gleichwahrscheinlich ausgewählt. Für Auswertungen auf Personenebene ist damit designbedingt eine Gewichtung notwendig. In der Praxis ist die Anwendung dieses sog. „Transformations-Gewichts“ nicht unumstritten, weil - wie wir auch hier im folgenden noch sehen werden - das zur Gewichtung herangezogene Merkmal „Haushaltsgröße“ selbst nicht frei von Verzerrungen ist (vgl. Hartmann 1990, Hartmann/Schimpl-Neimanns 1992). Wenn Aussagen über die Qualität von realisierten Stichproben getroffen werden sollen, liegt es jedoch nahe, daß entsprechend den Vorgaben des Designs verfahren und mithin designgewichtet ausgewertet wird (vgl. allerdings Hartmann (1990: 10), der anders verfährt).

14) Eine Analyse des Teilnahmeverhaltens beim ALLBUS 1994 mit Individualdaten aus den Einwohnermelderegistern für Teilnehmer und Nichtteilnehmer zeigt, daß die Unterrepräsentierung der Frauen tatsächlich ein Nonresponse-Effekt ist (vgl. Koch 1997: 112). Die Teilnahmequote der Frauen lag in Westdeutschland mit 49,7 Prozent knapp 10 Prozentpunkte unter der der Männer mit 59,1 Prozent. Auch in Ostdeutschland war die Ausschöpfung bei Männern mit 59,2 Prozent höher als bei Frauen mit 53,7 Prozent. Diese geringere Differenz zwischen den Geschlechtern im Osten steht im Einklang mit der hier beobachtbaren geringeren Unterrepräsentierung von Frauen in Ostdeutschland.

15) Auch wenn die einzelnen Abweichungen zumeist nicht signifikant sind, bedeutet dies nicht, daß sie tatsächlich zufällig erfolgen. Nicht signifikante Abweichungen können zufällig sein, müssen es aber nicht. Es spricht stichprobentheoretisch nichts gegen ihre Interpretation, wenn es hierfür plausible Gründe - im vorliegenden Fall die Gleichförmigkeit der Abweichung, die mit den bekannten Regelmäßigkeiten des Teilnahmeverhaltens übereinstimmt - gibt. Wären die Abweichungen zufällig, dann träte die Konstellation 'sechs mal negativ' mit der Wahrscheinlichkeit von $0,5^6 = 0,015625$ auf.

16) Zu erwähnen ist, daß auch beim Merkmal Alter die vorliegenden Ergebnisse für den ALLBUS 1994 in ihrer Grundstruktur mit den Befunden auf Basis der Registerangaben (vgl. Koch 1997: 112) übereinstimmen.

17) Dies gilt erstaunlicherweise auch für die Media-Analyse, für die man nicht nur wegen der hohen Ausschöpfung, sondern auch aufgrund der höheren Fallzahlen und des dadurch niedrigeren Stichprobenfehlers eigentlich eine bessere Anpassung an den Mikrozensus als bei den anderen Umfragen erwarten würde.

18) Darauf hinzuweisen ist, daß die in der Tabelle berichteten Abweichungen vom Mikrozensus die tatsächliche Untererfassung niedriger Bildungsabschlüsse vermutlich sogar noch unterschätzen. Da im Mikrozensus 1993 die Fragen zum Bildungsabschluß nicht der Auskunftspflicht unterlagen, weisen diese einen Item-Nonresponse von ca. 10 bis 15 Prozent auf. Analysen zeigen, daß dieser Item-Nonresponse selektiv erfolgt: insbesondere Bildungsabschlüsse unterer Qualifikationsniveaus werden im Mikrozensus unterschätzt (vgl. Riede/Emmerling 1994).

19) Hauptgrund hierfür ist die unterschiedliche Erfassung des Abschlusses „Polytechnische Oberschule“ der ehemaligen DDR.

20) Es kann nicht völlig ausgeschlossen werden, daß für dieses Ergebnis neben einem Nonresponse-Effekt auch Unterschiede im Antwortverhalten zwischen Mikrozensus einerseits und den Umfragen andererseits eine Rolle spielen. Im Mikrozensus wurde bei der betreffenden Frage die Antwortkategorie der Selbständigen nicht weiter differenziert, in allen sechs Umfragen dagegen wurde explizit zwischen selbständigen Landwirten, (akademischen) freien Berufen und sonstigen Selbständigen unterschieden (vgl. die Rekodierungen im Anhang). Da die Abfrage in West- und Ostdeutschland in gleicher Weise erfolgte, müßte eine Erklärung auf der Basis unterschiedlichen Antwortverhaltens allerdings annehmen, daß eine differenzierte Abfrage lediglich in Ostdeutschland zu systematisch anderen Einstufungen der Befragten führt als eine summarische.

21) Lediglich beim Sozialwissenschaften-Bus II 1993 stimmt der Anteil der Ledigen in Ostdeutschland nahezu exakt mit dem Ergebnis des Mikrozensus überein.

22) Erwähnenswert ist in diesem Zusammenhang, daß auch bei einer internen Validierung der Stichprobenqualität durch Sodeur (1997: 80ff.) der auf Basis einer

Einwohnermelderegister-Stichprobe durchgeführte ALLBUS 1994 sehr gut abschneidet. In den nach dem ADM-Verfahren realisierten ALLBUS-Erhebungen der Jahre 1988 bis 1992 (mit zum Teil weit höherer berichteter Ausschöpfung, Anm. Achim Koch) stellt Sodeur dagegen Abweichungen von einer reinen Wahrscheinlichkeitsauswahl fest, die nach seiner Vermutung mit dem größeren Spielraum der Interviewer bei der Auswahl der Befragten bei diesem Verfahren zusammenhängen.

23) Das Muster der im Gegenzug überrepräsentierten Haushaltsgrößen ist unterschiedlich. Während beim ALLBUS 1992 und Wohlfahrtssurvey 1993 in West- und Ostdeutschland Personen aus Haushalten mit drei und mehr Personen zu stark vertreten sind, ist bei den Umfragen des Sozialwissenschaften-Bus und der Media-Analyse im Westen die Überrepräsentierung von Personen aus Zwei- und Vierpersonenhaushalten besonders auffällig, im Osten die Überrepräsentierung von Personen aus Zwei- sowie Fünf(undmehr)personenhaushalten.

24) Ob dies auch für andere, hier nicht untersuchte Merkmale zutrifft, kann nicht beurteilt werden. Wie in der Einleitung ausgeführt wurde, ist der Nonresponse-Bias eine merkmalspezifische Größe. Verallgemeinerungen der Ergebnisse über die betrachteten Merkmale hinaus sind damit weder möglich noch beabsichtigt.

25) Für die Gestaltung der Feldarbeit ist es deshalb notwendig, über Mittel und Wege nachzudenken, wie man nicht nur die Ausschöpfung erhöhen, sondern gleichzeitig auch die Struktur der Stichprobe verbessern kann. Dies ist nur durch den *gezielten* Einsatz ausschöpfungssteigernder Maßnahmen (wie zusätzlicher callbacks, incentives, etc.) möglich, die dazu beitragen, bestehende Ungleichgewichte in den Teilnahmewahrscheinlichkeiten verschiedener Bevölkerungsgruppen abzubauen.

26) Ähnlich äußert sich Schnell (1997: 58f.). Auch in einer „offiziellen“ Darstellung des ADM-Designs wird darauf hingewiesen, daß zumindest beim ADM-Design ohne Adressenvorlauf die Gefahr besteht, daß Interviewer „unbequeme“ Adressen übergehen (vgl. von der Heyde/Löffler 1993: 52). In den Routinekontrollen der Umfrageinstitute können solche Abweichungen vom Stichprobenplan durch die Befragung einer falschen Person - im Unterschied etwa zur Totalfälschung eines Interviews - nicht entdeckt werden.

Erwähnenswert ist in diesem Zusammenhang im übrigen, daß insbesondere das in Deutschland bei face-to-face-Umfragen übliche Entlohnungssystem keinen Anreiz für die Interviewer bietet, sich mit „schwierigen“ Fällen zu befassen: honoriert werden lediglich realisierte Interviews, und zwar zu einem festen „Stückpreis“.

27) Nach Sommer (1987) wäre es im übrigen falsch, geschönte Ausschöpfungsergebnisse einseitig den Interviewern anzulasten. Nach seiner Ansicht liegt der Fehler im System, da Abweichungen von den Random-Anweisungen von den Umfrageinstituten stillschweigend erwartet und toleriert würden.

Literatur

- Arbeitsgemeinschaft Media-Analyse, o. J.: Dokumentation MA 94. Frankfurt/M.
- Braun, M./Eilinghoff, C./Gabler, S./Wiedenbeck, M., 1993: Methodenbericht zur „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1992. ZUMA-Arbeitsbericht 93/01.
- Burton, J., 1996: Public Attitudes and Responses to Survey Research: A Review of the Literature. Unpubl. paper. University of Essex.
- Esser, H./Grohmann, H./Müller, W./Schäffer, K.-A., 1989: Mikrozensus im Wandel. Untersuchungen und Empfehlungen zur inhaltlichen und methodischen Gestaltung. Bericht des wissenschaftlichen Beirates für Mikrozensus und Volkszählung. Frankfurt/M.
- Gabler, S./Hoffmeyer-Zlotnik, J. H. P. (Hrsg.), 1997: Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.
- GFM-GETAS, o. J.: Sozialwissenschaften-Bus II/III 1993 - Methodenbericht zur technischen Organisation und Durchführung. Hamburg.
- Goyder, J., 1987: The Silent Minority. Nonrespondents on Sample Surveys. Cambridge: Polity Press.
- Groves, R. M., 1989: Survey Errors and Survey Costs. New York: John Wiley & Sons.
- Groves, R. M./Couper, M. P., 1992: Correlates of Nonresponse in Personal Visit Surveys. S. 102-111 in: Proceedings of the American Statistical Association, Survey Research Methods Section.
- Hartmann, P. H., 1990: Wie repräsentativ sind Bevölkerungsumfragen? Ein Vergleich des ALLBUS und des Mikrozensus. ZUMA-Nachrichten 26: 7-30.
- Hartmann, P. H./Schimpl-Neimanns, B., 1992: Sind Sozialstrukturanalysen mit Umfragedaten möglich? Analysen zur Repräsentativität einer Sozialforschungsumfrage. Kölner Zeitschrift für Soziologie und Sozialpsychologie 44: 315-340.
- Infratest Sozialforschung, 1993: Wohlfahrtssurvey 1993 - Methodenbericht. München.
- Koch, A., 1995: Gefälschte Interviews: Ergebnisse der Interviewerkontrolle beim ALLBUS 1994. ZUMA-Nachrichten 36: 89-105.
- Koch, A., 1997: Teilnahmeverhalten beim ALLBUS 1994. Soziodemographische Determinanten von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft. Kölner Zeitschrift für Soziologie und Sozialpsychologie 49: 98-122.
- Koch, A./Gabler, S./Braun, M., 1994: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1994. ZUMA-Arbeitsbericht 94/11.
- Landgrebe, K. P., 1992: Ausschöpfungen. Planung und Analyse 19(2): 19-22.

- Lüttinger, P./Riede, Th., 1997: Der Mikrozensus: Amtliche Daten für die Sozialforschung. ZUMA-Nachrichten 41: 19-43.
- Porst, R., 1996: Ausschöpfungen bei sozialwissenschaftlichen Umfragen. Die Sicht der Institute. ZUMA-Arbeitsbericht 96/07.
- Riede, Th./Emmerling, D., 1994: Analysen zur Freiwilligkeit der Auskunftserteilung im Mikrozensus. Sind Stichprobenergebnisse bei freiwilliger Auskunftserteilung verzerrt? *Wirtschaft und Statistik* 9: 733-742.
- Schnell, R., 1997: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske + Budrich.
- Sodeur, W., 1997: Interne Kriterien zur Beurteilung von Wahrscheinlichkeitsauswahlen. *ZA-Information* 41: 58-82.
- Sommer, R., 1987: Der Mythos der Ausschöpfung. *Planung und Analyse* 14: 300-301.
- von der Heyde, Ch./Löffler, U., 1993: Die ADM-Stichprobe. *Planung und Analyse* 20(5): 49-53.

Anhang: Definition der Kategorien

1. Schulabschluss

In den Auswertungen für Ostdeutschland wurden Kategorie 1 und 2 zusammengefaßt.

Mikrozensus

1. Bis Volks-/Hauptschulabschluss:	-kein Schulabschluss -Haupt- oder Volksschulabschluss
2. Mittlere Reife, Fachhochschulreife:	-Realschulabschluss (Mittlere Reife) -polytechnische Oberschule der DDR -Fachhochschulreife
3. Abitur, Fachhochschul-/Hochschulabschluss	-Hochschulreife (Abitur) -Fachhochschule, Ingenieurschule -Hochschule

ALLBUS, Wohlfahrtssurvey

1. Bis Volks-/Hauptschulabschluss:	-Schule beendet ohne Abschluß -Volks-/Hauptschulabschluss, Polytechnische Oberschule mit Abschluß 8. oder 9. Klasse
2. Mittlere Reife, Fachhochschulreife:	-Mittlere Reife, Realschulabschluss, Polytechnische Oberschule mit Abschluß 10. Klasse -Fachhochschulreife (Abschl. einer Fachobersch. etc.)
3. Abitur, Fachhochschul-/Hochschulabschluss	-Abitur, Erweiterte Oberschule mit Abschluß 12. Klasse (Hochschulreife) -Fachhochschulabschluss (auch Abschluß einer Ingenieurschule) -Hochschulabschluss

Schüler und Personen mit „anderem Schulabschluss“ sind in der Prozentuierung nicht berücksichtigt.

Sozialwissenschaften-Bus

1. Bis Volks-/Hauptschulabschluss:	-von der Schule abgegangen ohne Abschluß, vor Erreichen einer 9., 10., oder 11. Klasse -Volks-, Hauptschulabschluss
2. Mittlere Reife, Fachhochschulreife:	-Mittlere Reife, Realschulabschluss, Fachschulreife -Polytechnische Oberschule mit 10. Klasse Abschluß (vor 1965: 8. Klasse Abschluß) -Fachhochschulreife, Fachgebundene Hochschulreife, Abschluß einer Fachoberschule
3. Abitur, Fachhochschul-/Hochschulabschluss	-Abitur, allgemeine Hochschulreife, Erweiterte Oberschule (EOS)

-Fachhochschulabschluß (auch Ingenieurschulabschluß)
-Hochschulabschluß ohne Diplom
-Hochschulabschluß mit Diplom/Promotion

Schüler und Personen mit „anderem Schulabschluß“ sind in der Prozentuierung nicht berücksichtigt.

Media-Analyse

1. Bis Volks-/Hauptschulabschluß:	-Volksschule/Grundschule, Hauptschule
2. Mittlere Reife, Fachhochschulreife:	-Mittlere Reife bzw. weiterführende Schule ohne Abitur
3. Abitur, Fachhochschul-/Hochschulabschluß:	-Abitur, Hochschulreife -Studium (Universität, Hochschule, Fachhochschule, Akademie, Polytechnikum)

2. Stellung im Erwerbsleben

Mikrozensus

1. Hauptberuflich erwerbstätig:	-Personen, deren normalerweise geleistete Arbeitszeit je Woche 17 oder mehr Stunden beträgt
2. Nicht hauptberuflich erwerbstätig:	-Nicht erwerbstätige Personen -Personen, deren normalerweise geleistete Arbeitszeit je Woche 16 oder weniger Stunden beträgt

ALLBUS

1. Hauptberuflich erwerbstätig:	-Hauptberufliche Erwerbstätigkeit, ganztags -Hauptberufliche Erwerbstätigkeit, halbtags -In Kurzarbeit
2. Nicht hauptberuflich erwerbstätig:	-Nicht erwerbstätig -Nebenher erwerbstätig

Wohlfahrtssurvey

1. Hauptberuflich erwerbstätig:	-Hauptberuflich erwerbst., ganztags (35 Std. und mehr) -Hauptberuflich erwerbst., halbtags (unter 35 Std.)
2. Nicht hauptberuflich erwerbstätig:	-Nebenher erwerbstätig, regelmäßig (unter 20 Std.) -Nebenher erwerbstätig, unregelmäßig -Nicht erwerbstätig (auch arbeitslos) -Wehr- oder Zivildienst

Sozialwissenschaften-Bus

1. Hauptberuflich erwerbstätig:	-Vollzeit erwerbstätig, Wochen-Arbeitszeit West: 35 Std. und mehr, Ost: 40 Std. und mehr -Teilzeit erwerbstätig, Wochen-Arbeitszeit West: 17 bis 34 Std., Ost: 20 bis 39 Std.
---------------------------------	--

	-Auszubildende/Lehrlinge
2. Nicht hauptberuflich erwerbstätig:	-Nicht erwerbstätig -Teilzeit oder stundenweise erwerbstätig, Wochen- Arbeitszeit West: unter 17 Std., Ost: unter 20 Std. -(Vollzeit-) Schüler

Media-Analyse

1. Hauptberuflich erwerbstätig:	-Voll berufstätig -Teilweise berufstätig -In Ausbildung: Lehrling
2. Nicht hauptberuflich erwerbstätig:	-Vorübergehend arbeitslos -Rentner, Pensionär (nicht mehr voll berufstätig) -In Ausbildung: Schüler, Student -Nicht berufstätig

3. Berufliche Stellung**Mikrozensus**

1. Arbeiter:	-Arbeiter, Heimarbeiter
2. Angestellte:	-Angestellte, Zivildienstleistende
3. Beamte:	-Beamte, Richter, Polizei, BGS, Soldaten
4. Selbständige:	-Selbständige, mit und ohne Beschäftigte -Mithelfende Familienangehörige
5. In Ausbildung:	-Kaufmännisch-technische Auszubildende -Gewerbliche Auszubildende

ALLBUS

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellter
3. Beamte:	-Beamter/Richter/Berufssoldat
4. Selbständige:	-Selbständiger Landwirt (auch Genossenschaftsbauer) -Akademischer freier Beruf -Selbständiger in Handel, Gewerbe, Industrie, Dienstleistung u.a. -Mithelfender Familienangehöriger
5. In Ausbildung:	-In Ausbildung

Wohlfahrtssurvey

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellte
3. Beamte:	-Beamte/Richter/Berufssoldaten
4. Selbständige:	-Selbständige Landwirte -Akademische freie Berufe -Sonstige Selbständige (außerhalb der Landwirtschaft) -Mithelfende Familienangehörige
5. In Ausbildung:	-In Ausbildung

Sozialwissenschaften-Bus

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellter
3. Beamte:	-Beamter/Richter/Berufssoldat
4. Selbständige:	-Selbständiger Landwirt/Selbständiger Bauer oder Gärtner (auch Genossenschaftsbauer) -Akademischer freier Beruf -Selbständiger in Handel, Gewerbe, Industrie, Dienstleistung u.a. (auch PGH) -Mithelfende Familienangehörige
5. In Ausbildung:	-In Ausbildung

Media-Analyse

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellte
3. Beamte:	-Beamte
4. Selbständige:	-Selbständige -Freie Berufe -Selbständige Landwirte (Vollerwerbsbetrieb)
5. In Ausbildung:	-In Ausbildung; Lehrling

ANALYSEMÖGLICHKEITEN DES MIKROZENSUS

BERNHARD SCHIMPL-NEIMANN

Seit Anfang 1997 können die faktisch anonymisierten Einzeldaten des Mikrozensus 1995 von der Forschung relativ einfach vom Statistischen Bundesamt bezogen werden. Im Vergleich zu Umfragedaten aus der empirischen Sozialforschung liegen die Vorteile des Mikrozensus in der Qualität und Größe der Stichprobe. Die Nutzer des Mikrozensus sind jedoch auch mit Einschränkungen konfrontiert, die daraus resultieren, daß das Fragenprogramm vorwiegend am Datenbedarf von Verwaltung und Politik orientiert ist. Dieser Beitrag wird anhand von Beispielen die Auswertungsmöglichkeiten des Mikrozensus praxisbezogen und methodenkritisch darstellen. Die Fragestellungen konzentrieren sich auf die Bereiche Arbeitsmarkt, Haushalte, Familien und Sozialstruktur.

Since the beginning of 1997, access to the Mikrozensus 1995 has been improved greatly for social scientists. In comparison to social surveys, official microdata offers certain advantages, such as a high quality sample and a large sample size. Nevertheless, users of the Mikrozensus face also some limitations which are linked to the fact that the Mikrozensus is largely generated within a framework determined by administrative and official needs. The paper briefly reviews the potential and shortcomings of using the Mikrozensus in empirical social research. The emphasis is on methods. For labour market, household, family, and social structure, the research potential is discussed in more detail.

1. Einleitung

Nach der Vereinbarung zwischen dem Statistischen Bundesamt und dem Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF), die faktisch anonymisierten Einzeldaten des Mikrozensus 1995 für die Forschung als scientific use file zur Verfügung zu stellen (siehe ZUMA-Nachrichten 40, S. 96-98), wird der Mikrozensus inzwischen von einer Reihe von einzelnen Wissenschaftlern und Forschungseinrichtungen für die Bearbeitung verschiedenster Forschungsfragen genutzt.

Mit einem Stichprobenumfang von einem Prozent der Personen und Haushalte ist der Mikrozensus die größte Mehrthemenumfrage der Bundesrepublik Deutschland und enthält vielfältige Informationen über die wirtschaftliche und soziale Lage der Bevölkerung. Der Forschung steht eine 70%-Substichprobe des Mikrozensus 1995 mit rund 512.000 Personen in 224.000 Haushalten zur Verfügung. Zu folgenden Themenbereichen liegen in rund 200 Variablen Angaben vor: soziodemographische Merkmale, Staatsangehörigkeit und Zusatzangaben für Ausländer, Familien- und Haushaltszusammenhang, Erwerbsbeteiligung und berufliche Verhältnisse, Schulbesuch und Aus- und Weiterbildung, Quellen des Lebensunterhalts, Einkommen und Kranken- und Rentenversicherung, Fragen zur Gesundheit, Behinderteneigenschaft und betriebliche Altersvorsorge. Ausführliche Datenbeschreibungen sind im World Wide Web unter der Adresse <http://www.zuma-mannheim.de/data/microdata> zu finden. In den ZUMA-Nachrichten 41 haben Lüttinger/Riede (1997) den Mikrozensus ausführlich vorgestellt und im Überblick Nutzungsmöglichkeiten genannt.

Der große Stichprobenumfang des Mikrozensus ist eine für die Forschung wichtige Eigenschaft, da dadurch differenzierte Analysen auch kleinerer Bevölkerungsgruppen möglich sind.¹⁾ Neben der Stichprobengröße ist die hohe Ausschöpfung des Mikrozensus von Vorteil. Aufgrund der Auskunftspflicht liegt die Antwortquote der Haushalte bei rund 97 Prozent. Mit Mikrozensusdaten können somit Verteilungs- und Zusammenhangsmaße für die Population verlässlich geschätzt werden. Weiterhin kann der Mikrozensus für die normalen Bevölkerungsumfragen, die vielfach Ausschöpfungsquoten unter 60 Prozent haben, auch als Referenzstatistik dienen.

Diese Vorteile, die sich der Forschung mit der Analyse des Mikrozensus bieten, gelten unter sehr allgemeinen Annahmen. Genauere Aussagen über die Qualität, mit der soziale Tatbestände mit den Daten des Mikrozensus abgebildet werden, können erst getroffen werden, wenn weitere Faktoren bekannt sind, insbesondere über die in allen Befragungen mehr oder weniger vorhandenen systematischen Fehler, wie z.B. Antwortfehler oder Verzerrungen aufgrund von Item-Nonresponse.

Bei der Verwendung von Daten des Mikrozensus gelten im Prinzip dieselben Vor- und Nachteile, die auch für die Sekundäranalyse anderer Daten gelten (vgl. Alba et al. 1994a). Ganz allgemein ist festzuhalten, daß die Sekundäranalyse von Daten Replikationen und Reanalysen erlaubt und damit den wissenschaftlichen Diskussionsprozeß fördert. Des weiteren trägt die Mehrfachnutzung von Daten zu einem effizienteren Mitteleinsatz bei. Dies gilt besonders in bezug auf amtliche Mikrodaten, deren Analysemöglichkeiten durch die amtliche Statistik alleine bei weitem nicht ausgeschöpft werden.

Die wichtigste Einschränkung bei der Sekundäranalyse besteht darin, daß die Daten nicht immer den Anforderungen entsprechen, die für die Bearbeitung einer Fragestellung notwendig sind. Der Forscher kann nur von den vorhandenen Variablen ausgehen und muß eine möglichst gute Beziehung dieser Angaben zu seiner Ausgangsfragestellung bzw. den zu testenden Hypothesen herstellen. Die hierbei entstehenden Probleme der Operationalisierung und Indikatorenbildung führen teilweise zur Änderung der Ausgangsfragestellung. Diese Einschränkung gilt mehr oder weniger für alle Sekundäranalysen. Während jedoch die Daten, die von der Forschung selbst erhoben werden, für Forschungszwecke konzipiert sind, ist im Fall des Mikrozensus zu beachten, daß das Erhebungsprogramm primär am Datenbedarf von Verwaltung und Politik orientiert ist; unter anderem werden keine Fragen zu Einstellungen, Handlungsabsichten oder Handlungsdeutungen gestellt.

Der Beitrag des Mikrozensus zur Bearbeitung von Forschungsfragen kann nur vor dem Hintergrund der jeweiligen Fragestellung bewertet werden. Aufgrund der Eigenschaft des Mikrozensus als Mehrthemenumfrage der amtlichen Statistik ist es nicht möglich, an dieser Stelle einen Gesamtüberblick (vgl. dazu Esser et al. 1989) zu geben. Es werden nur einige Beispiele dargestellt, die entsprechend der Hauptaufgabe des Mikrozensus, als „laufende Repräsentativstatistik über die Bevölkerung und den Arbeitsmarkt“ über den Stand und die Entwicklung der Sozialstruktur, die Gründung und Veränderung von Haushalten und Familien sowie besonders über das Arbeitsmarktgeschehen statistische Informationen zu liefern, ausgewählt wurden.

2. Arbeitsmarkt und Erwerbstätigkeit

Im Zentrum der Arbeitsmarktforschung, die mit der gestiegenen Verfügbarkeit von Individualdaten mehr empirische Anwendungsmöglichkeiten gefunden hat, stehen unter anderem Fragen des Rückgangs der Erwerbsbeteiligung insbesondere in Ostdeutschland, nach den Faktoren von Langzeitarbeitslosigkeit, zur Erwerbsarbeit von Frauen in Verbindung mit Problemen der Vereinbarkeit von Berufstätigkeit und Mutterschaft, nach Veränderungen der Tätigkeits- und Qualifikationsstruktur von Erwerbstätigen sowie nach Zusammenhängen zwischen Ausbildung und Entlohnung. Übersicht 1 gibt einen Überblick darüber, welche Variablen im anonymisierten Mikrozensusfile 1995 für die Bearbeitung dieser und weiterer Fragestellungen verwendet werden können.

Abgrenzung des Erwerbsstatus

Um die oben genannten Fragestellungen zu untersuchen, ist eine Unterscheidung der Bevölkerung nach ihrer Beteiligung am Erwerbsleben notwendig. Für die Erfassung des Erwerbsstatus einer Person können subjektive oder objektive Kriterien verwendet werden.

Übersicht 1: Verfügbare Variablen für die Arbeitsmarktforschung*

- Art des überwiegenden Lebensunterhalts
- Beteiligung am Erwerbsleben
- Bei Erwerbstätigen (bei Nichterwerbstätigen: zur früheren Tätigkeit):
 - Regelmäßige oder gelegentliche Tätigkeit; geringfügige Beschäftigung
 - Arbeitszeit; Voll-, Teilzeitarbeit; tatsächliche und normalerweise geleistete Arbeitszeit
 - Grund für Teilzeittätigkeit
 - Art, Befristung des Arbeitsvertrages
 - Stellung im Beruf (frühere); Stellung im Betrieb
 - Wirtschaftszweig (früherer)
 - Beschäftigung im öffentlichen Dienst (frühere)
 - Ausgeübter Beruf (früherer)
 - Berufs-, Betriebswechsel
 - Suche und Grund der Suche nach anderer Tätigkeit
 - Zweite Erwerbstätigkeit (Stellung im Beruf; Wirtschaftszweig; tatsächliche und normalerweise geleistete Arbeitszeit)
- Bei Erwerbslosen (Arbeitslose und Arbeitssuchende)
 - Bezug von Arbeitslosengeld, -hilfe
 - Anlaß und Dauer der Arbeitssuche; Methode der Arbeitssuche
 - Art und Umfang der gesuchten Tätigkeit
 - Verfügbarkeit für eine neue Arbeitsstelle; Gründe für die Nichtverfügbarkeit
- Bei Nichterwerbstätigen (siehe auch bei Erwerbstätigen):
 - frühere Erwerbstätigkeit
 - Zeitpunkt und Beendigungsgrund für letzte Tätigkeit
- daneben eine Vielzahl kombinierbarer Variablen; z.B.: demographische Merkmale, Haushalts- und Familienzusammenhang, Krankenversicherung, Rentenversicherung, Aus- und Weiterbildung, Einkommen, Ausländerfragen etc.
- Fragen zur sozialen oder privaten Pflegeversicherung und zur Pflegebedürftigkeit (ab 1996)

* Zusammengestellt nach Krug et al. 1994: 263-264.

Beispielsweise ordnen sich im ALLBUS die Befragten hinsichtlich ihres hauptsächlichen Erwerbsstatus selbst einer Kategorie zu (Selbsteinschätzung). Im Unterschied

dazu folgt der Mikrozensus bei der Abgrenzung von Erwerbstätigen und Nichterwerbstätigen dem Labour-Force-Konzept der International Labour Organization (ILO). Nach dieser Definition gilt jede Person im Alter von 15 und mehr Jahren als Erwerbstätiger, die in der Berichtswoche (i.d.R. die letzte April-Woche) mindestens eine Stunde gegen Entgelt oder als Selbständiger bzw. als Mithelfender Familienangehöriger gearbeitet hat. Als Erwerbs- bzw. Arbeitslose gelten nach dem ILO-Konzept Personen, die nicht nach obiger Definition erwerbstätig sind, für eine Arbeitsaufnahme innerhalb eines bestimmten Zeitraums (sofort bzw. innerhalb von zwei Wochen) zur Verfügung stehen und innerhalb der letzten vier Wochen aktive Schritte der Arbeitssuche unternommen haben (Recktenwald 1997; Rinne 1994: 169ff.; Berié/Mayer 1983). Um die Abgrenzung von Erwerbstätigen, Erwerbslosen und Nichterwerbstätigen umzusetzen, wird im Mikrozensus eine Reihe von Fragen verwendet, damit auch jene Befragten korrekt eingestuft werden, die sich „subjektiv“ nicht als Erwerbstätige einschätzen, aber „objektiv“ nach den ILO-Standards dazu gehören.

Erst ab dem Mikrozensus 1996 weist das Statistische Bundesamt auch Ergebnisse aus, die hinsichtlich des Verfügbarkeits-Kriteriums der engen ILO-Definition von Erwerbslosen folgen. Im Mikrozensus 1995 wurde bei der Abgrenzung der Erwerbslosen die Verfügbarkeit noch nicht berücksichtigt. Es ist an dieser Stelle deshalb wichtig darauf hinzuweisen, daß im Grundfile neben der Variablen Erwerbsstatus nicht nur alle für diese Typisierung verwendeten Ausgangsvariablen enthalten sind, sondern auch die Angaben zur Verfügbarkeit für eine Arbeitsstelle.²⁾ Damit können Nutzer den ILO-Erwerbsstatus selbst umsetzen, was vor allem hinsichtlich internationaler Vergleiche hilfreich ist. Die Gegenüberstellung der i.d.R. vom Statistischen Bundesamt verwendeten Definition und des ILO-Konzeptes des Erwerbsstatus in Tabelle 1 zeigt, daß die Berücksichtigung des Verfügbarkeitskriteriums zu einer Verringerung des Anteils der Erwerbslosen führt.³⁾ Da bekanntermaßen Arbeitslosenquoten häufig politisch kontrovers diskutiert werden, ist in diesem Zusammenhang besonders zu beachten, wie die statistischen Ergebnisse von der verwendeten Definition abhängen.

Tabelle 1: Verteilung der Bevölkerung am Ort der Hauptwohnung nach dem Erwerbsstatus (in Prozent) und nach Berücksichtigung der Verfügbarkeit für eine neue Arbeitsstelle

Erwerbsstatus	ohne Berücksichtigung der Verfügbarkeit für eine neue Arbeitsstelle	mit Berücksichtigung der Verfügbarkeit für eine neue Arbeitsstelle
Erwerbstätige	44,1	44,1

Erwerbslose bzw. ILO-Erwerbslose	4,9	4,0
Nichterwerbspersonen	51,0	51,9
Zusammen (n=815.549)	100	100
Erwerbslosenquote*	10,1	8,2

Quelle: Mikrozensus 1995 (ZUMA-File); Bevölkerung am Ort der Hauptwohnung; hochgerechnete/gewichtete Fallzahlen.

*) Anteil der Erwerbslosen an den Erwerbspersonen (Erwerbstätige und Erwerbslose).

Anzumerken ist des weiteren, daß die mit dem Mikrozensus ermittelte Erwerbslosenquote aus konzeptionellen Gründen nicht mit der auf Basis der Arbeitsmarktstatistik der Bundesanstalt für Arbeit berechneten Arbeitslosenquote vergleichbar ist. Einerseits werden in der Arbeitsmarktstatistik nur registrierte Arbeitslose gezählt, während die Meldung beim Arbeitsamt nach dem im Mikrozensus verwendeten ILO-Konzept keine Rolle spielt. Andererseits können auch registrierte Arbeitslose in begrenztem Umfang erwerbstätig sein; solche Fälle werden im Mikrozensus nach der ILO-Definition aber als Erwerbstätige gezählt (Gruber 1995; Rinne 1994: 199).

Erwerbstätige und Erwerbslose können nach einer Reihe weiterer Merkmale gegliedert werden (Übersicht 1). In Abbildung 1 und 2 werden diese beiden Gruppen nach ausgewählten Charakteristika dargestellt, die sich an aktuellen Arbeitsmarktthemen orientieren.

Abbildung 1 zeigt, daß in den neuen Bundesländern der Anteil von Erwerbstätigen mit befristeten Arbeitsverträgen rund doppelt so hoch ist wie in Westdeutschland, wobei in Ostdeutschland Frauen häufiger als Männer in einem befristeten Arbeitsverhältnis stehen. In welchem Ausmaß ostdeutsche Erwerbstätige – gezwungenermaßen aufgrund der Strukturveränderungen seit der Vereinigung – mobil waren, zeigt der im Vergleich zu Westdeutschland wesentlich höhere Anteil von Betriebs- und Berufswechseln sowie die häufigere Teilnahme an Maßnahmen der beruflichen Fortbildung und Umschulung; letzteres trifft besonders für Frauen in den neuen Bundesländern zu.

In Abbildung 2 wird die ILO-Definition von Erwerbslosen (Verfügbarkeit für den Arbeitsmarkt) zugrunde gelegt. Bei den Verteilungsunterschieden nach dem Bezug von Arbeitslosengeld bzw. -hilfe sowie einer gesuchten Voll- oder Teilzeittätigkeit ist ein markanter Unterschied zwischen den Frauen in West- und Ostdeutschland feststellbar. Während ostdeutsche Frauen in der Verteilung dieser Merkmale ähnliche Muster wie die Männer aufweisen, was auf eine ähnliche Erwerbsorientierung hinweist, beziehen westdeutsche Frauen weit seltener Arbeitslosengeld oder -hilfe, sie suchen seltener eine Vollzeittätigkeit und statt dessen häufiger eine Teilzeittätigkeit als Männer oder Frauen in Ostdeutschland. Vierzig Prozent der ostdeutschen erwerbslosen Frauen sind zwei

Jahre oder länger auf Arbeitssuche und rund sechzig Prozent haben als Beendigungsgrund ihrer letzten Tätigkeit Entlassung angegeben, was im Zusammenhang mit den oben bereits angesprochenen Berufs- und Betriebswechselln der Erwerbstätigen in Ostdeutschland zu sehen ist. Für tiefergehende Analysen struktureller Aspekte der Erwerbs- bzw. Arbeitslosigkeit stehen im Datensatz neben den in Abbildung 2 genannten Variablen weitere Informationen zur Verfügung (s. Übersicht 1).

Abbildung 1: Ausgewählte Charakteristika von Erwerbstätigen nach Geschlecht und Region (in Prozent bezogen auf jeweils alle Erwerbstätigen)

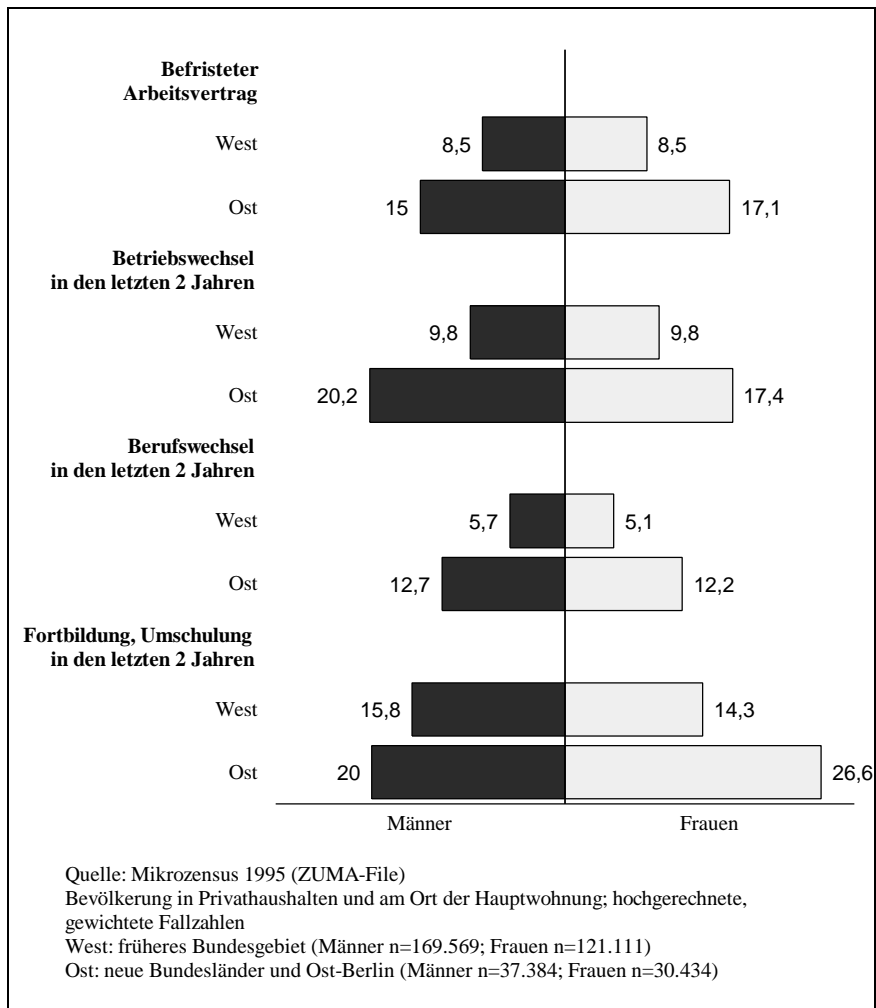
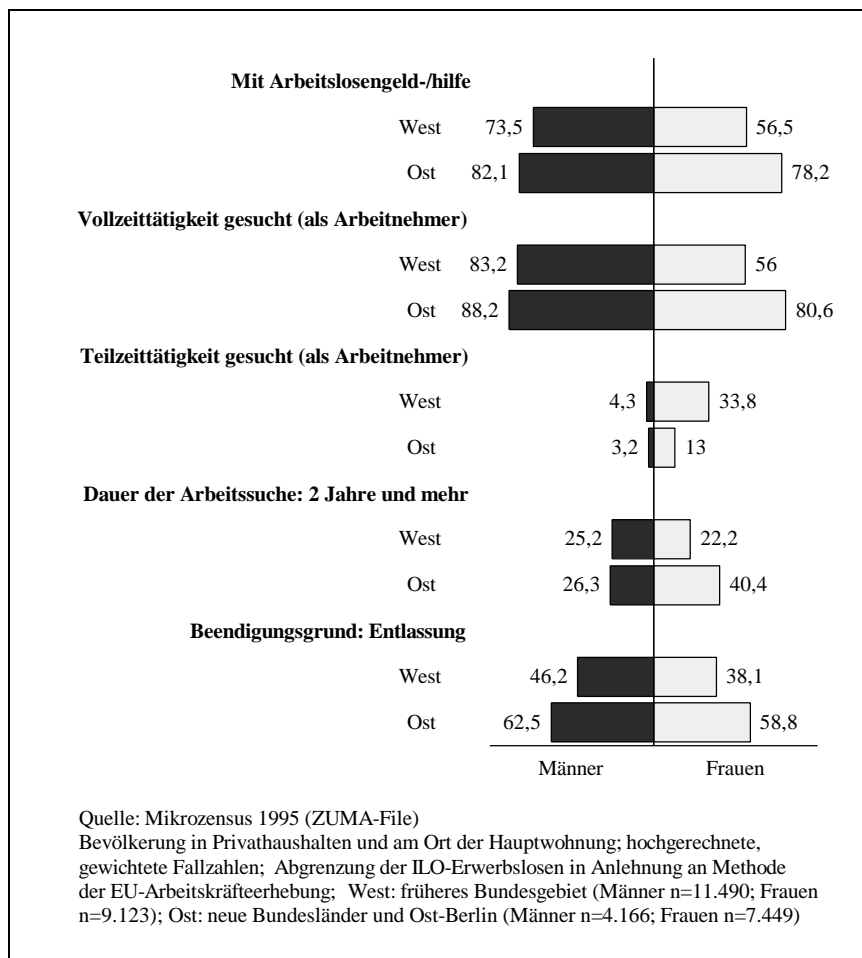


Abbildung 2: Ausgewählte Charakteristika von Erwerbslosen (ILO) nach Geschlecht und Region (in Prozent bezogen auf jeweils alle Erwerbslosen)



Seit 1990 wird im Mikrozensus eine eigene Frage nach der *Ausübung einer geringfügigen Beschäftigung* gestellt, die zum Kranz der sogenannten Leitfragen zur Ermittlung bzw. Typisierung des Erwerbsstatus gehört.⁴⁾ Es ist davon auszugehen, daß

der Mikrozensus den Umfang geringfügiger Beschäftigung unterschätzt (vgl. DIW 1997; Statistisches Bundesamt 1997; Kohler et al. 1996). Dies liegt im wesentlichen an dem Berichtszeitraum, auf den sich die Ergebnisse zur Erwerbstätigkeit beziehen. Ende April, der Berichtswoche des Mikrozensus, treten bestimmte geringfügige Beschäftigungen seltener auf, wie beispielweise saisonabhängige Beschäftigungen und Ferienjobs von Schülern und Studenten. Zum Teil dürfte die im Vergleich zu anderen Erhebungen niedrige Zahl geringfügig Beschäftigter im Mikrozensus auch mit den Schwierigkeiten zusammenhängen, die komplexe sozialrechtliche Definition geringfügiger Beschäftigung im Rahmen einer Befragung umzusetzen. Zwar wird direkt nach der Ausübung einer geringfügigen Beschäftigung gefragt, es kann aber vermutet werden, daß Befragte, die Gelegenheitstätigkeiten z.B. als Haushaltshilfen ausüben, diese Tätigkeit nicht immer als Erwerbstätigkeit verstehen.

Die obigen Schaubilder können nur knappe Hinweise auf das Auswertungspotential und die vielfältigen Kombinationsmöglichkeiten geben. Im folgenden werden unter spezifischen inhaltlichen Aspekten weitere Möglichkeiten und Einschränkungen der Sekundäranalyse des Mikrozensus angesprochen.

Berufliche Mobilität

Als Querschnittsdatensatz enthält der Mikrozensus 1995 mit den Informationen über Berufs- und Betriebswechsel seit 1993 nur begrenzte Angaben über die berufliche Mobilität. Zur früheren Erwerbstätigkeit von Nichterwerbstätigen sind Angaben darüber enthalten, wann und mit welchem Grund diese Tätigkeit beendet wurde. Im Mikrozensus wird nur der ausgeübte, nicht jedoch auch der erlernte Beruf erfragt, so daß beispielsweise Berufswechsel nicht analysiert werden können.

In der EU-Arbeitskräfteerhebung, die mit einem durchschnittlichen Auswahlsatz von 0,45 Prozent im Rahmen des Mikrozensus durchgeführt wird, werden Retrospektivangaben erhoben, die für die Analyse beruflicher Mobilität sehr interessant sind. Es liegen für alle Personen dieser Teilstichprobe Angaben zur Situation vor einem Jahr (Erwerbstätigkeit, Stellung im Beruf, Wirtschaftszweig, Wohnsitzwechsel) vor; für Erwerbstätige wird die Zahl der Personen in der Betriebseinheit und der Zeitpunkt des Arbeitsbeginns erfragt, für Arbeitslose/-suchende der Zeitpunkt des letzten Kontakts mit dem Arbeitsamt, der Wunsch nach einer Erwerbstätigkeit und die Situation unmittelbar vor der Arbeitssuche. Aus rechtlichen Gründen können diese speziellen Mikrozensusvariablen der EU-Arbeitskräfteerhebung erst mit Inkrafttreten des Mikrozensusgesetzes 1996, d.h. ab dem Mikrozensus 1996, an die Forschung weitergegeben werden.

Einkommen

Die Einkommensfrage zählt in Umfragen zu den „heiklen“ Fragen und ist in normalen Umfragen häufig mit einem hohen Item-Nonresponse behaftet. Dagegen liegen diese Antwortausfälle im Mikrozensus bei rund vier Prozent.

Die Einkommensangaben im Mikrozensus haben vorwiegend die Funktion eines ergänzenden Merkmals zur Beschreibung der sozio-ökonomischen Lagen. Erfragt wird das Nettoeinkommen im Monat April aus allen Einkommensquellen. Dazu gehören Einkommen aus selbständiger Tätigkeit (1/12 des Jahreseinkommens), Löhne und Gehälter, sowie wohlfahrtsstaatliche Leistungen (Arbeitslosengeld, bzw. -hilfe, Kindergeld, Renten etc.), private Unterstützungen, Einkommen aus Vermietung, Verpachtung und Kapitalvermögen sowie Sachbezüge. Bei selbständigen Landwirten und allen Mithelfenden Familienangehörigen entfällt die Einkommensangabe. Die Angaben werden gruppiert in 18 Kategorien mit nach oben offener Randklasse (7.500 und mehr DM) erhoben.

Für die mikroökonomische Untersuchung von Aspekten der Lohnstruktur und der Erwerbseinkommen, wie sie etwa mit Regressionen des Mincer-Typs⁵⁾ vorgenommen wird, ist der Mikrozensus mit der Einschränkung verwendbar, daß nicht das (Brutto- oder Netto-) Erwerbseinkommen vorliegt, sondern das Nettoeinkommen aus allen Einkommensquellen. Unter Verwendung der Angaben zur Frage nach der überwiegenden Quelle des Lebensunterhalts ist es möglich, die Analysen auf jene Befragte zu konzentrieren, die ihren Lebensunterhalt überwiegend aus Erwerbstätigkeit bestreiten.

Probleme bei Proxi-Angaben am Beispiel des Einkommens

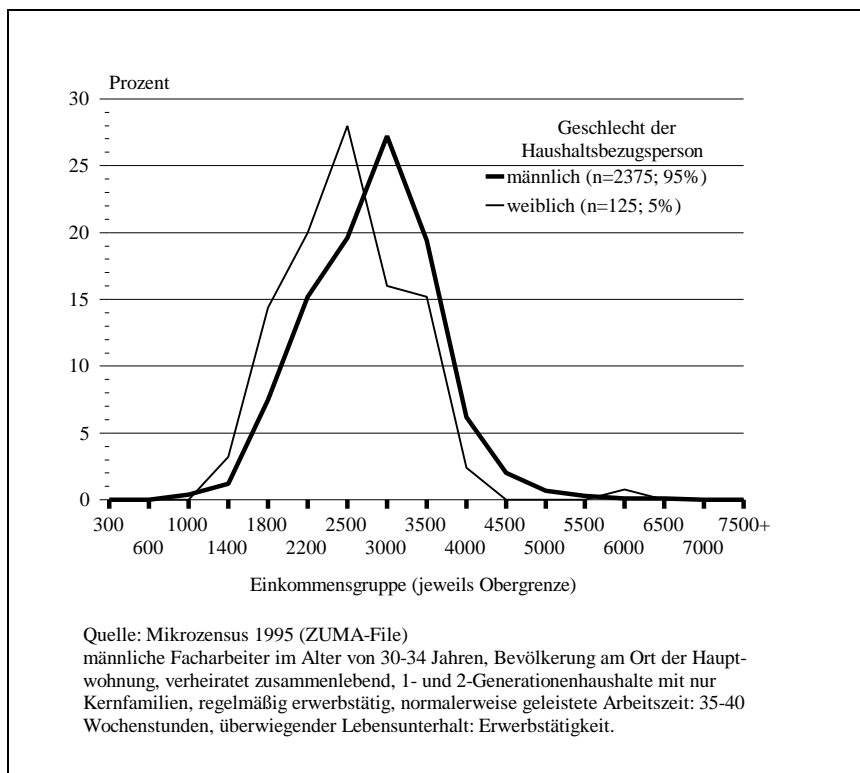
Am Beispiel der Einkommensvariablen soll an dieser Stelle auf das im Mikrozensus vorliegende Problem der Proxi-Interviews hingewiesen werden; d.h. wenn Angaben nicht direkt von den Personen selbst, sondern von Dritten gemacht werden. Aus Kostengründen wird beim Mikrozensus darauf verzichtet, bei Erwachsenen ausschließlich Selbstauskünfte zu erfassen. Trifft der Interviewer nicht alle Personen im Haushalt bei der Befragung an, können die Auskünfte über die nicht anwesenden Personen auch von anderen Personen im Haushalt gegeben werden.

Es ist bekannt, daß es bei Proxi-Interviews häufiger als bei Selbstauskünften zu Antwortausfällen kommt und die Antwortqualität geringer ist (Dawe/Knight 1997; Blair et al. 1991). Diese variiert mit der Schwierigkeit der Frage und dem Grad der über die anderen Personen benötigten Information. Während Angaben zum Alter und zum

Familienstand unproblematisch sind, besitzt nicht jeder genaue Kenntnisse zur korrekten Beantwortung der Fragen nach Beruf und Branche etc.

Eine direkte Untersuchung der Validität von Fremdauskünften ist mit dem Mikrozensus nicht möglich, da keine Angaben über die auskunftgebenden Personen vorliegen. Um dennoch ansatzweise dieses Problem zu beleuchten, kann als Indikator für die Auskunftsperson die Angabe zur Haushaltsbezugsperson (siehe dazu Abschnitt 3, Typisierungen) genutzt werden.

Abbildung 3: Verteilung der persönlichen Nettoeinkommen von Facharbeitern nach Geschlecht der Haushaltsbezugsperson



In Abbildung 3 werden für die nach Homogenitätskriterien abgegrenzte Gruppe männlicher Facharbeiter im Alter von 30-34 Jahren die Einkommen dargestellt. Aufgrund der

für dieses Beispiel getroffenen Auswahl von 1- und 2-Generationenhaushalten kann davon

ausgegangen werden, daß die Facharbeiter überwiegend selbst Auskunft gegeben haben, wenn das Geschlecht der Haushaltsbezugsperson männlich ist; dies ist mit 95 Prozent die überwiegende Mehrheit der ausgewählten Gruppe. Ist dagegen das Geschlecht der Haushaltsbezugsperson weiblich, kann vermutet werden, daß häufiger die Ehefrauen über die Einkommen der Männer Auskunft gegeben haben.⁶⁾ Wie zu sehen, weichen beide Verteilungen voneinander ab und offensichtlich unterschätzen Frauen die Einkommen der Männer. Diese hier aus Platzgründen nur exemplarisch dargestellten Verteilungsdifferenzen zeigen sich auch bei anderen Gruppen.⁷⁾

Ausländer

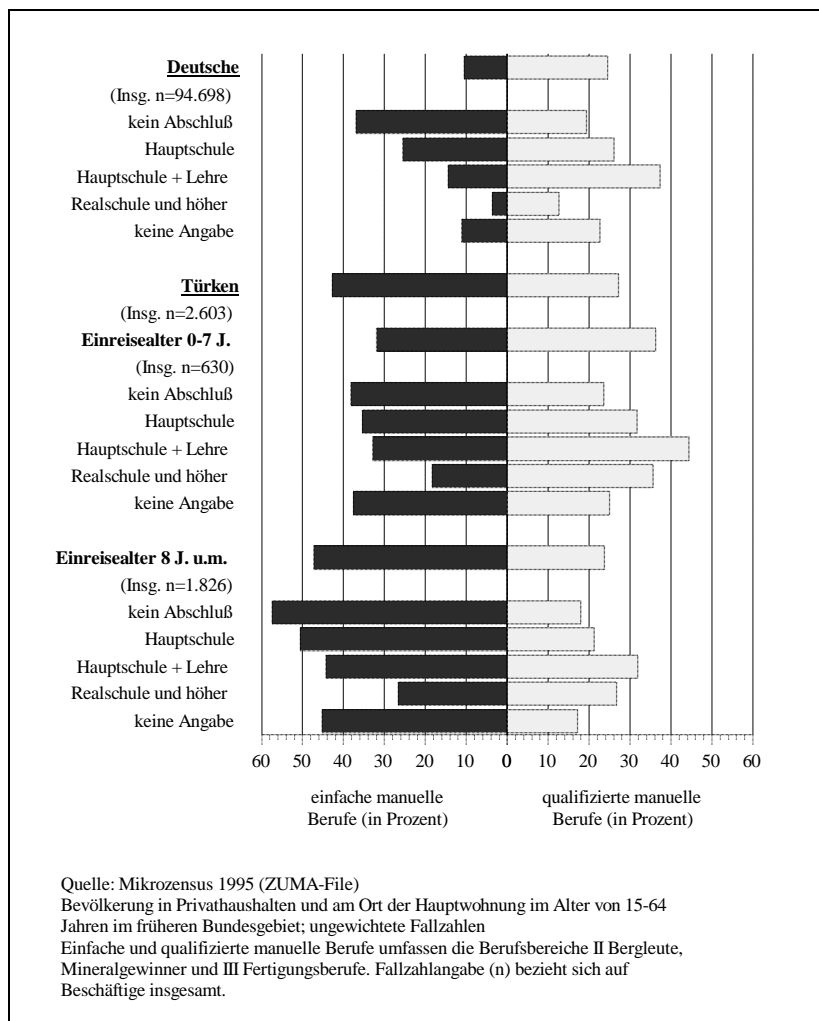
Für Analysen zur Ausländerbeschäftigung stehen im anonymisierten Grundfile detaillierte Informationen über die Staatsangehörigkeit von 10 Nationalitäten sowie 15 Nationalitätengruppen zur Verfügung. Aus Anonymisierungsgründen sind Kategorien kleinerer Ausländergruppen, die in der Population unter 50.000 Einwohner umfassen, zusammengefaßt, z.B. Tschechien, Slowakien und Ungarn.

In der Forschung zur Integration von Ausländern bzw. zur Arbeitsmarktsegregation dienen häufig Angaben über die Rückkehrabsichten von Migranten als Indikator für den Stand der Integration. Diese Informationen liegen zwar im Mikrozensus nicht vor, es ist aber möglich, ersatzweise auf andere Indikatoren der familiären Bindung zum Heimatland zurückzugreifen. Im Mikrozensus werden an Ausländer Fragen nach im Heimatland lebenden Ehegatten, Eltern oder Kinder gestellt. Diese Angaben sind von der Auskunftspflicht befreit und weisen deshalb Antwortausfälle von etwa 20 Prozent auf (Emmerling/Riede 1994; Riede/Emmerling 1994). Angaben zum Zuzugsjahr können zur Abgrenzung von Migranten-Generationen genutzt werden. Bei Verwendung dieser Variablen ist neben den Antwortausfällen aufgrund freiwilliger Auskünfte zu beachten, daß bei der Unterbrechung des Aufenthalts um 6 Monate und mehr das Jahr der (jüngsten) Rückkehr erfaßt wird. Da das Geburtsland nicht erfragt wird, können deshalb bei der Ermittlung von Migrantengenerationen Klassifikationsfehler auftreten.⁸⁾

In Abbildung 4 wird zum Themenbereich ethnischer Arbeitsmarktsegregation ein einfaches Auswertungsbeispiel gegeben. In den alten Bundesländern sind 57 Prozent aller männlichen Ausländer in Fertigungsberufen oder als Bergleute beschäftigt. Dagegen üben nur 35 Prozent aller deutschen männlichen abhängig Erwerbstätigen einen Beruf dieser Gruppe aus. Unterteilt man diese Berufe noch weiter in einfache und qualifizierte manuelle Berufe, sind 10 Prozent der Deutschen in einfachen und 24 Prozent in qualifizierten manuellen Berufen beschäftigt. Von allen türkischen Beschäftigten – der zahlen-

mäßig stärksten Ausländergruppe – sind dagegen 42 Prozent in einfachen und 27 Prozent in qualifizierten manuellen Berufen beschäftigt.⁹⁾

Abbildung 4: Deutsche und türkische männliche Beschäftigte in einfachen und qualifizierten manuellen Berufen



Zur Klärung der Frage, inwiefern ein längerer Aufenthalt zur Verbesserung der Berufschancen von Ausländern - hier der Türken - beiträgt, wird mit Hilfe des Zuzugsjahres das Einreisealter ermittelt und bei sieben Jahren dichotomisiert. In der ersten Gruppe werden in Deutschland geborene und bis zum Alter von sieben Jahren eingereiste Personen zusammengefaßt. Dem liegt die Annahme zugrunde, daß das längere Aufwachsen im Gastland und die Eingliederung in die deutsche Regelschule mit größeren Assimilationseffekten und besseren Berufschancen einhergeht als bei der zweiten Gruppe; 95 Fälle, für die keine Angabe zum Einreisealter vorliegen, bleiben außer Betracht. Während in Deutschland geborene oder bis zum Alter von sieben Jahren eingereiste Türken zu 32 Prozent einen einfachen manuellen Beruf ausüben, trifft dies für 47 Prozent der zweiten Gruppe zu. Da die Ausbildungsqualifikationen die beruflichen Chancen beeinflussen, dient die höchste erreichte allgemeine und berufliche Ausbildung als Kontrollvariable. In Abbildung 4 sind der Übersichtlichkeit halber nur die Anteile von Befragten in einfachen und qualifizierten Berufen abgetragen; alle anderen Berufe bilden die Differenz zu 100 Prozent. Für Deutsche zeigt sich ein starker Zusammenhang zwischen Bildungsqualifikation und Berufsgruppe. Mit steigender Qualifikation sinkt der Anteil von Personen in einfachen manuellen Berufen. Der Zusammenhang zwischen Bildungsqualifikation und Berufsgruppe ist bei den Türken ebenfalls vorhanden, aber wesentlich schwächer als bei den Deutschen. Von den türkischen abhängig Erwerbstätigen mit Hauptschulabschluß üben in der ersten Gruppe der bis zum Alter von sieben Jahren Eingereisten zu 35 Prozent einen einfachen manuellen Beruf aus; in der zweiten Kategorie gehören 51 Prozent zu dieser Berufsgruppe. Ein zusätzlicher Ausbildungsabschluß reduziert den Anteil einfacher manueller Berufe in der ersten Gruppe auf 33 Prozent und in der zweiten Gruppe auf 44 Prozent. Das Einreisealter stellt sich damit in diesem Beispiel als der im Vergleich zur Bildungsqualifikation wichtigere Faktor heraus.

Frauenerwerbstätigkeit

Mikrozensusdaten enthalten detaillierte Informationen zur Arbeits- und Berufswelt und sind somit sehr gut zur Untersuchung der Struktur der Frauenerwerbstätigkeit, zu geschlechtsspezifischen Disparitäten und für segmentations- und segregationstheoretische Fragestellungen geeignet. Berufe und Wirtschaftszweige liegen zum Beispiel im Grundfile sehr differenziert als Dreisteller mit 366 bzw. 214 Ausprägungen vor.

Bei Analysen zur Frauenerwerbstätigkeit werden zur Schätzung von Funktionen des Arbeitsangebots bzw. der Erwerbsbeteiligung üblicherweise neben den sozio-ökonomi-

schen Faktoren wie Alter, Qualifikation und Einkommen etc. auch Angaben über die Zahl und das Alter von Kindern sowie das Haushaltseinkommen oder das Einkommen des Ehegatten berücksichtigt. Solche Angaben des Familien-/Haushaltskontextes können im Mikrozensus leicht ermittelt werden. Während bei der Schätzung des Arbeitsangebots von Frauen auf der Basis kleinerer Stichproben aus Fallzahlgründen oft nur zusammengefaßte Angaben über die Kinder (z.B. Zahl der Kinder unter 6 Jahren usw.) verwendet werden können, läßt sich mit dem Mikrozensus der Zusammenhang wesentlich differenzierter untersuchen, da auch kleinere Altersabgrenzungen noch ausreichend besetzt sind. Dies erlaubt eine Überprüfung der funktionalen Form der geschätzten Beziehung zwischen Erwerbstätigkeit und der Zahl von Kindern in verschiedenen Altersstufen. Nachteilig ist jedoch, daß nur Informationen über jene Kinder zur Verfügung stehen, die noch im elterlichen Haushalt leben.

3. Haushalts- und Familienstruktur

Wie bereits angesprochen, können mit dem Mikrozensus Auswertungen sowohl auf Personen- als auch auf Familien- und Haushaltsebene vorgenommen werden, da die anonymisierte Substichprobe alle Personen eines Haushalts enthält und diese Auswertungseinheiten eindeutige Identifikatoren besitzen. Darüber hinaus enthält der Datensatz eine Reihe von sogenannten Bandsatzerweiterungen zum Haushalt und zur Familie (z.B. Zahl der Erwerbstätigen), zur Bezugsperson des Haushalts (z.B. Alter, Beruf) und verschiedene Typisierungen (z.B. Haushaltstyp nach Verwandtschaftsbeziehungen).

Für haushalts- und familienspezifische Auswertungen ist die Definition dieser Einheiten von zentraler Bedeutung und wird deshalb kurz genannt. Haushalte bestehen aus Personen, die allein oder zusammen mit anderen eine wirtschaftliche Einheit bilden, d.h. gemeinsames Wohnen und Wirtschaften ist eine notwendige Voraussetzung. Die amtliche Statistik definiert Familie in Anlehnung an Empfehlungen der Vereinten Nationen. Als Familien zählen Ehepaare mit und ohne Kinder sowie alleinerziehende ledige, verheiratet getrenntlebende, geschiedene und verwitwete Väter und Mütter mit ihren im gleichen Haushalt lebenden ledigen Kindern. Bei der Befragung zählen ledige Stief-, Adoptiv- oder Pflegekinder wie leibliche Kinder und können deshalb im Datensatz von diesen nicht unterschieden werden.

Im Mikrozensus 1995 wurde keine direkte Frage zum Haushalt und zur Familie gestellt (z.B. zum Haushaltseinkommen).¹⁰⁾ Als Beispiel für familienspezifische Auswertungsmöglichkeiten sind in Tabelle 2 die Nettoeinkommen von deutschen Familien mit zwei Kindern im Alter bis zu 15 Jahren nach dem Erwerbsstatus der Eltern und der Region (alte/neue Bundesländer) ermittelt worden. Die Ergebnisse zeigen die deutlichen Unter-

schiede der materiellen Lagen in West und Ost. In ostdeutschen Familien sind erheblich häufiger als im Westen beide Elternteile erwerbstätig, wobei die Erwerbstätigkeit von Müttern wesentlich zum Familieneinkommen beiträgt.

Tabelle 2: Median des Familiennettoeinkommens von Familien verheiratet zusammenlebender Eltern mit 2 Kindern im Alter bis zu 15 Jahren und deutscher Familienbezugsperson nach Region und ILO-Erwerbsstatus der Eltern

ILO-Erwerbsstatus der Eltern		früheres Bundesgebiet		neue Bundesländer	
Vater	Mutter	Familien %	Median DM	Familien %	Median DM
erwerbstätig	erwerbstätig	46,5	4670	73,3	3980
erwerbstätig	nicht erwerbstätig	49,5	3730	20,7	2820
erwerbslos	erwerbstätig	0,7	2850	2,6	3150
erwerbslos	nicht erwerbstätig	1,2	1970	1,7	1640
nicht erwerbstätig	zusammen *	2,0	2720	1,7	2800
Zusammen		10155	4090	2875	3640
		100%		100%	

Quelle: Mikrozensus 1995 (ZUMA-File)

Bevölkerung am Familienwohnsitz; 2-Generationenhaushalte mit nur geradlinig verwandten Personen; 4-Personen-Familien; ungewichtete Fallzahlen.

Definition des ILO-Erwerbsstatus der Eltern in Anlehnung an Methode der EU-Arbeitskräfteerhebung.

Für die Berechnung des Familieneinkommens wurden die jeweiligen Klassenmitten verwendet; bei der nach oben offenen Randklasse „7.500 DM und mehr“ wurde die Klassenmitte auf 10.000 DM festgesetzt.

Ohne Familien, für die keine Einkommensangabe vorliegt (Familien mit mithelfenden Familienangehörigen, selbständigen Landwirten, ohne Angabe; n=990(West), 165 (Ost)).

*) Wegen kleiner Fallzahlen in dieser Gruppe sind Erwerbstätige, Erwerbslose und nicht Erwerbstätige zusammengefaßt.

Typisierungen

Typisierungen des Haushalts und der Familie basieren auf den Angaben zum Familienzusammenhang, welche ausschließlich die Verwandtschaftsverhältnisse zur Haushaltsbezugsperson abbilden. Bezugsperson des Haushalts ist die erste im Fragebogen des Haushalts eingetragene Person, wobei die Reihenfolge beim Interview vorgegeben ist (Ehegatten, Kinder, Verwandte, Familienfremde) und eine Bezugsperson mindestens 15 Jahre alt sein muß. Diese erhebungstechnische Definition der Haushaltsbezugsperson ist nicht gleichzusetzen mit dem Haushaltsvorstand. Als Haushaltsvorstand gilt i.d.R. jene Person, welche die sozio-ökonomische Lage des

Haushalts am stärksten bestimmt („Hauptverdiener“). Haushaltsvorstände können mit den anonymisierten Einzeldaten unter Verwendung der Einkommensvariablen und der Haushaltsidentifikatoren leicht nachträglich vom Forscher selbst abgegrenzt werden.

Koresidenzprinzip

Im Mikrozensus werden nur Angaben zu Personen erhoben, die im ausgewählten Haushalt leben (zu diesem Koresidenzprinzip vgl. Höhn 1988). Aus rechtlichen Gründen werden keine Fragen zu Personen gestellt, die nicht bzw. nicht mehr im Haushalt leben. Diese Personen können mit einem Haushalt noch verbunden sein und dessen soziale und ökonomische Lage beeinflussen. Dies betrifft z.B. Studenten, die nicht mehr (auch nicht mit Zweitwohnsitz) bei den Eltern wohnen, aber noch von ihren Eltern abhängig sind. Es fehlen deshalb auch, wie bereits erwähnt, Angaben zur Zahl der von einer Frau geborenen Kinder oder Angaben zur sozialen Herkunft. Allerdings kann für Kinder, die noch im elterlichen Haushalt leben, der Status der Eltern problemlos ermittelt werden. Einschränkend ist jedoch zu beachten, daß aufgrund des Auszugs aus dem Elternhaus z.B. nur noch 78 Prozent aller 20jährigen Jugendlichen bei ihren Eltern leben, während dies für 98 Prozent der 15jährigen zutrifft.

Das Koresidenzprinzip, die Gleichstellung von Stief-/Adoptiv- und Pflegekindern mit leiblichen Kindern sowie die ausschließliche Erfassung der Verwandtschaftszusammenhänge zur Haushaltsbezugsperson haben für die empirische Haushalts- und Familienforschung gewisse Einschränkungen bei der Untersuchung neuer Formen des Zusammenlebens zur Folge, die von der „Normalfamilie“ abweichen. Es sei deshalb darauf hingewiesen, daß im Datensatz mit den Variablen 'Nicht verwandte Personen im Haushalt' und 'Haushaltstyp: Alleinstehend/Alleinerziehende' Informationen vorliegen, die eine Analyse untypischer Formen des Zusammenlebens erlauben, zum Beispiel von Nichteheleichen Lebensgemeinschaften.¹¹⁾

4. Sozialstrukturanalyse

Schulbesuch, Bildungsabschluß

Der Mikrozensus enthält bildungsstatistische Angaben (Schulbesuch, Bildungsabschluß, berufliche Weiterbildung), die für die Sozialstrukturanalyse, aber auch für die Arbeitsmarktforschung von großem Interesse sind.

Im Mikrozensus wird seit 1991 der Schulbesuch bzw. die Bildungsbeteiligung nach der Internationalen Standardklassifikation für das Bildungswesen (ISCED) mit folgenden

Kategorien erhoben: Kindergarten /-krippe /-hort; Klassenstufe 1-4 / 5-10 / 11-13 einer allgemeinbildenden Schule; Berufliche Schule; Fachhochschule; Hochschule. Damit wird zwar die internationale Vergleichbarkeit der Ergebnisse verbessert und das frühere Zuordnungsproblem von Schülern integrierter Gesamtschulen zu den Schulformen des dreigliedrigen Schulsystems gelöst, aber mit dem Wegfall der Schulformen – eine Besonderheit des (west-) deutschen Bildungssystems – ist zum Beispiel die Analyse schichtspezifischer Schulbesuchsraten nicht mehr möglich, da Haupt- und Realschüler sowie Gymnasiasten bis einschließlich der Klassenstufe 10 nicht unterscheidbar sind. Jedoch können Besucher der gymnasialen Oberstufe (Klassenstufe 11-13) nach sozialer Herkunft, Geschlecht, Staatsangehörigkeit etc. differenziert werden. Hinsichtlich der Ermittlung der sozialen Herkunft gilt die mit dem Koresidenzprinzip verbundene Einschränkung (s.o.). Angaben über kaufmännische/technische und gewerbliche Auszubildende können mit Hilfe der Variablen 'Stellung im Beruf' ermittelt werden.

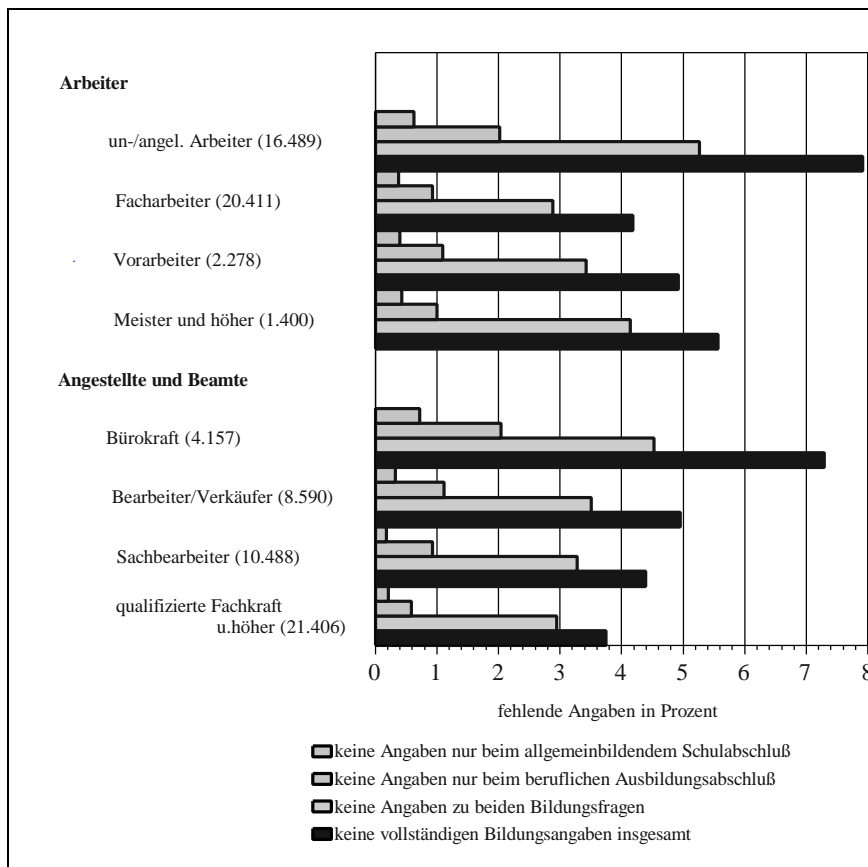
Für Personen, die keine allgemeinbildende Schule mehr besuchen, liegen folgende Informationen vor: Angaben zum Abschluß einer allgemeinbildenden Schule, Art des erreichten höchsten allgemeinen Schulabschlusses sowie letzter beruflicher Ausbildungsabschluß. Für Absolventen einer Fachhochschule und einer Hochschule wird außerdem die Hauptfachrichtung des letzten Abschlusses erfragt. Diese Variablen können zur Konstruktion einer Bildungsskala, z.B. durch Kombination allgemeiner und beruflicher Abschlüsse, oder für die Rekodierung typischer Ausbildungsdauern verwendet werden. Neben den Abschlüssen werden Informationen über die Teilnahme an einer beruflichen Fortbildung in den letzten zwei Jahren nach dem Ausbildungsort und der Ausbildungsdauer erhoben.

Aufgrund der unterschiedlichen Struktur der Bildungssysteme in der BRD und DDR werden bestimmte DDR-spezifische Abschlüsse in speziellen Kategorien erfaßt (vgl. Schimpl-Neimanns/Lüttinger 1993). Beim allgemeinbildenden Schulabschluß ist die „Polytechnische Oberschule“ als eigene Kategorie angegeben, dagegen werden Absolventen der „Erweiterten Oberschule“ und des Bildungsgangs „Berufsausbildung mit Abitur“ in der Kategorie Abitur zusammengefaßt. Hinsichtlich des beruflichen Ausbildungsabschlusses gilt die Gleichstellung der „Beruflichen Teilausbildung“ mit dem Abschluß einer Lehre. Der Besuch einer Fachschule der DDR wird hingegen als eigene Antwortkategorie erfaßt, jedoch ohne Differenzierungsmöglichkeit nach der Hauptfachrichtung.

Für die meisten Fragen des Mikrozensus besteht für die befragten Haushalte Auskunftspflicht. Ausgenommen davon wurden in den Mikrozensusen 1991-1995 jedoch unter anderem die Bildungsfragen, bei denen der Item-Nonresponse rund acht Prozent beträgt.

Untersuchungen des Statistischen Bundesamtes haben gezeigt, daß Bildungsangaben des Mikrozensus infolge von systematischen Antwortausfällen zu verzerrten Bildungsverteilungen führen und damit die Ergebnisqualität beeinträchtigt ist (Emmerling/Riede 1994; Riede/Emmerling 1994). Mit dem Mikrozensusgesetz 1996 wurde teilweise die Konsequenz aus diesen Ergebnissen gezogen, indem die Bildungsfragen für Personen bis zum 50. Lebensjahr seit 1996 wieder auskunftspflichtig sind.

Abbildung 5: Nonresponse bei Bildungsfragen des Mikrozensus 1995



Quelle: Mikrozensus 1995 (ZUMA-File); 18-64jährige abhängig erwerbstätige Männer; früheres Bundesgebiet; Bevölkerung am Ort der Hauptwohnung; ungewichtete Fallzahlen (n).

In Anlehnung an Riede/Emmerling (1994) wird in Abbildung 5 für abhängig Erwerbstätige in Westdeutschland dargestellt, wie die Höhe des Nonresponse bei verschiedenen Gruppen variiert. Als Proxi-Variable für die Qualifikation dient hierbei die Kombination der beruflichen und betrieblichen Stellung. Bei den Angestellten und Beamten ist ein inverser Zusammenhang zwischen Nichtbeantwortung und betrieblicher Stellung festzustellen; d.h. je niedriger die mittels betrieblicher Stellung gemessene Qualifikation, desto häufiger werden die Bildungsfragen nicht beantwortet. Diese Beziehung gilt bei den Arbeitern nur für die un- und angelernten Arbeiter sowie für Facharbeiter, während der Item-Nonresponse bei Vorarbeitern und Meistern im Vergleich zu Facharbeitern höher ist.

Diese Abhängigkeit der Nichtbeantwortung der Bildungsfragen vom Qualifikationsniveau bedeutet, daß die gültigen Bildungsangaben nicht ohne weitere Korrekturen auf die Population der Bevölkerung hochgerechnet werden können. Nicht angemessen wäre auch eine Verteilung der fehlenden Werte proportional zu den gültigen Bildungsangaben. Für Möglichkeiten der Korrektur von Nonresponse kann an dieser Stelle nur auf die weiterführende Literatur verwiesen werden (Little/Schenker 1995; Winship/Mare 1989; Berk 1983).

Berufsklassifikationen

Zur Untersuchung der Schichtung der Gesellschaft und zur Darstellung von sozialen Klassenlagen werden in der Regel auf der Basis von Berufsklassifikationen und weiteren Merkmalen Schicht-, Klassen- oder Prestigeskalen gebildet. Im Mikrozensus liegen viele Angaben vor, die für die Konstruktion solcher Variablen verwendet werden können: Beruf, Wirtschaftszweig, Stellung im Beruf, Stellung im Betrieb, überwiegend ausgeübte Tätigkeit und Lage des Arbeitsplatzes. Der Forscher ist damit nicht auf die grobe Gliederung nach sozialrechtlichen Kategorien angewiesen (Selbständige, Beamte, Angestellte, Arbeiter), die häufig bei der Verwendung veröffentlichter Mikrozensusergebnisse benutzt wird. Bei der Arbeit mit den anonymisierten Einzeldaten erweisen sich insbesondere in bezug auf die differenzierte Abbildung auch kleiner Gruppen, wie z.B. Selbständigen in Freien Berufen oder leitenden Angestellten, die großen Fallzahlen des Mikrozensus als wichtiger Vorteil für die Sozialforschung; für Beispiele kann hier nur auf verschiedene Arbeiten hingewiesen werden (siehe Wirth/Lüttinger 1998; Handl 1996; Alba et al. 1994b; Schimpl-Neimanns/Lüttinger 1993). An dieser Stelle soll auf zwei Einschränkungen in bezug auf die Variablen Stellung im Betrieb und Beruf hingewiesen werden.

Für die Erfassung vertikaler Aspekte der Schichtung bzw. der sozialen Klassenlage sind die Angaben zur Stellung im Betrieb von zentraler Bedeutung. Für Selbständige wird unterschieden nach der Zahl der Beschäftigten (0-4, 5 und mehr); für abhängig Beschäftigte und mithelfende Familienangehörige liegen sieben verschiedene Kategorien vor, die nach der Funktion innerhalb des Betriebes und nach der Autonomie gegliedert sind und von 'Bürokräft/Schreibkraft/Angelernter Arbeiter/Nicht-Facharbeiter' bis 'Direktor/Amtsleiter/ Geschäftsführer/Betriebsleiter/Werksleiter' reichen. Die Frage zur Stellung im Betrieb richtet sich nur an Erwerbstätige. Für Nichterwerbstätige können hilfsweise und im wesentlichen nur Angaben zur (früheren) Stellung im Beruf und zum (früheren) Beruf für die Schicht- oder Klassenzuordnung herangezogen werden.¹²⁾

Die meisten in sozialwissenschaftlichen Erhebungen verwendeten Schicht- oder Klassenskalen basieren auf der Internationalen Standardklassifikation der Berufe (ISCO), welche auch internationale Vergleiche ermöglicht. Der Mikrozensus verwendet dagegen die Klassifizierung der Berufe (Ausgabe 1992). Eine Umsetzung und Verwendung evtl. vorhandener Rekodierungsschemata auf den Mikrozensus setzt somit die Vergleichbarkeit der Klassifizierung der Berufe mit der ISCO-Klassifikation voraus. Diese Vergleichbarkeit ist zwar prinzipiell gegeben, im anonymisierten Grundfile aus technischen Gründen jedoch nicht herstellbar¹³⁾, so daß Schicht- oder Klassenskalen nur unter Verwendung der oben genannten Variablen konstruiert werden können.

5. Stichprobenfehler, Gewichtung

Stichprobenfehler

Bei der Datenanalyse wird in der Praxis bei der Verwendung von Standard-Statistikprogrammen von der Annahme einer einfachen Zufallsstichprobe ausgegangen. Der Mikrozensus gehört jedoch nicht zu den einfachen Zufallsstichproben, sondern weist als einstufige geschichtete Flächen- bzw. Klumpenstichprobe ein komplexes Design auf. Für die Hochrechnung von Stichprobenergebnissen auf die Grundgesamtheit muß der zufallsbedingte Stichprobenfehler berücksichtigt werden, der vom Auswahlplan abhängt. Im folgenden können nur die wichtigsten Grundprinzipien des komplexen Auswahlplans skizziert werden, welche die Schätzung des Stichprobenfehlers beeinflussen (ausführlicher informieren Krug et al. 1994: 239ff. und Meyer 1994).

In bezug auf die regionale und fachliche Schichtung im Mikrozensus kann von einer Verringerung des Stichprobenfehlers im Vergleich zu einer einfachen Zufallsstichprobe ausgegangen werden.¹⁴⁾ Hinsichtlich der Klumpung auf Zählbezirksebene gilt dies jedoch nicht. Die Auswahlbezirke bestehen aus benachbarten Gebäuden oder

Gebäudeteilen, wobei alle Haushalte eines ausgewählten Zählbezirks, und damit die in diesen Haushalten lebenden Personen, in die Stichprobe aufgenommen werden. Durchschnittlich gehören neun Wohnungen zu einem Zählbezirk. Im Vergleich zur einfachen Zufallsstichprobe führt die Klumpung im allgemeinen zu einer Vergrößerung des Stichprobenfehlers. Je ähnlicher die Merkmale innerhalb der Klumpen, je uneinheitlicher die Klumpengrößen und je größer die Klumpen sind, um so stärker fällt diese relative Vergrößerung des Stichprobenfehlers aus.

Aus Gründen der statistischen Geheimhaltung und des Datenschutzes enthalten die anonymisierten Einzeldaten, die an Forscher weitergegeben werden, grundsätzlich keine Informationen zum Stichprobenplan, d.h. keine Daten zum Zählbezirk und zu regionalen Schichten, die über die Bundesländer hinausgehen. Damit läßt sich einerseits die varianzverringende Schichtung in der Fehlerrechnung nicht berücksichtigen, andererseits kann die Intra-Klassenkorrelation aufgrund fehlender Klumpenidentifikation nicht ermittelt werden, so daß man davon ausgehen muß, daß bei der Varianzschätzung unter Nichtberücksichtigung des Design und unter Annahme einer einfachen Zufallsstichprobe der Stichprobenfehler unterschätzt wird.

Für die Hochrechnung einfacher statistischer Kennwerte, z.B. von Anteilswerten, können näherungsweise die vom Statistischen Bundesamt ermittelten Design-Effekt-Faktoren verwendet werden (siehe Statistisches Bundesamt 1996). Diese Faktoren geben für eine Zellenbesetzung bestimmter Merkmalskombinationen das durchschnittliche Verhältnis des Standardfehlers unter Berücksichtigung von Schichtung und Klumpung zum Standardfehler unter Annahme einer einfachen Zufallsauswahl wieder.

In bezug auf Zusammenhangsmaße und Anpassungsmaße, die mit multivariaten Verfahren ermittelt werden, können diese Design-Effekt-Faktoren nicht verwendet werden, so daß in der Forschungspraxis keine korrekte Fehlerrechnung unter Berücksichtigung des Designs möglich ist. Ersatzweise können Varianzschätzungen mittels Zufallsgruppen bzw. Resampling durchgeführt werden (Lee et al. 1989; Rendtel 1991; Rohwer 1994; Wolter 1985). Bisher durchgeführte Testanwendungen dieses Verfahrens bei Logit-Modellen zur Frauenerwerbstätigkeit und zur Bildungsungleichheit ergaben keine anderen statistischen Schlüsse als bei Annahme einer einfachen Zufallsstichprobe. Für eine Bewertung des Problems fehlender Designinformationen im Mikrozensus-Grundfile sind jedoch weitere Tests und die Verwendung alternativer Verfahren der Varianzschätzung nötig.

Gewichtung

Von den statistischen Ämtern werden die Mikrozensusergebnisse anhand von Eckzahlen der laufenden Bevölkerungfortschreibung hochgerechnet. Zu diesem Zweck wird ein Hochrechnungsfaktor auf der Basis von regionalen Eckzahlen zu sechs Anpassungsklassen gebildet, der im wesentlichen das Verhältnis der Sollzahlen aus der Bevölkerungfortschreibung zu den Istzahlen der im Mikrozensus Befragten abbildet.¹⁵⁾

Nutzer der faktisch anonymisierten Mikrozensusdaten können für die sogenannte freie Hochrechnung, die von den statistischen Ämtern für die Ergebnisdarstellung verwendet wird, auf den im Datensatz enthaltenen Hochrechnungsfaktor für Haushalte und Familien zurückgreifen.¹⁶⁾ Im Durchschnitt beträgt das Hochrechnungsgewicht 1,13; es liegt für Deutsche bei 1,1 und für Ausländer bei 1,5. Somit liegt im Mikrozensus im Vergleich zur Bevölkerungfortschreibung eine Untererfassung vor. Der insbesondere für Ausländer hohe Wert von 1,5 dürfte allerdings nur teilweise mit Befragungsproblemen des Mikrozensus in Verbindung stehen, sondern scheint zum größten Teil mit den Problemen der Bevölkerungfortschreibung, vor allem bezüglich nichterfaßter Abmeldungen bei den Melderegistern zusammenhängen.¹⁷⁾ Mit der Gewichtung werden die Mikrozensusergebnisse auf Ergebnisse der Bevölkerungfortschreibung adjustiert, was zur Übertragung von Fehlern der Bevölkerungfortschreibung auf den Mikrozensus führen kann.

Da der Hochrechnungsfaktor des Mikrozensus kein Designgewicht, sondern ein Korrekturgewicht ist, ist bei inferenzstatistischen Analysen eine Gewichtung nicht notwendig. Eine Gewichtung in diesem Zusammenhang ist auch nicht sinnvoll, weil dadurch die Zahl der Beobachtungen künstlich vergrößert und der Standardfehler verkleinert wird.

6. Schluß

Mit den verbesserten Nutzungsmöglichkeiten der anonymisierten Einzeldaten des Mikrozensus 1995 bieten sich vielfältige Möglichkeiten der Datenanalyse zur Bearbeitung verschiedenster Fragestellungen. Daß der Mikrozensus für die Sozialwissenschaften von größtem Interesse ist, liegt nicht nur am Umfang und in der Qualität der Stichprobe, sondern auch in der Vielfalt der mit den Daten untersuchbaren Themen. In diesem Beitrag wurden einige Möglichkeiten skizziert und einfache Beispiele vorgestellt.

Die beschriebenen Einschränkungen und Grenzen bei der Auswertung von Mikrozensusdaten können zusammenfassend folgenden Gründen zugeordnet werden. Wie in vielen anderen Mehrthemenumfragen stößt man auch im Mikrozensus bei der Bearbeitung von

spezifischen Forschungsfragen auf nicht erhobene oder nicht detailliert genug erfaßte Themen. Eine Reihe von Beschränkungen ergibt sich daraus, daß der Mikrozensus als Instrument der amtlichen Statistik nur auf der Basis entsprechender Rechtsgrundlagen durchgeführt werden kann. Seit Mitte der achtziger Jahre ist diesbezüglich eine verstärkte Verrechtlichung festzustellen, die sich unter anderem darin äußert, daß im Mikrozensusgesetz und in Verordnungen nicht nur die zulässigen Fragen aufgeführt, sondern auch die Antwortkategorien vorgegeben werden. Aus der Sicht des Gesetzgebers besteht die Hauptaufgabe des Mikrozensus darin, statistische Ergebnisse für Parlamente, Regierungen und Verwaltung in Bund und Ländern bereitzustellen. Aus diesem Gründen konzentriert sich das Fragenprogramm auf die Erhebung „objektiver“ Sachverhalte und es fehlen Fragen zu Einstellungen und Handlungsabsichten etc.

Inhaltliche Analyserestriktionen, die sich aus der Beschränkung in der Erfassung von Angaben der Personen eines Haushalts zum Erhebungszeitpunkt (Koresidenzprinzip) ergeben, wurden bei der Darstellung der Analysemöglichkeiten zum Thema Frauenerwerbstätigkeit und von Formen des Zusammenlebens angesprochen. Unter methodischen Aspekten wurden Probleme der Fremdauskunft (Proxi-Interviews), des Item-Nonresponse und der Varianzschätzung diskutiert. Zu diesen Methodenfragen ist noch weitere Forschung nötig.

Grundsätzlich können die inhaltlichen und methodischen Grenzen in der Sekundäranalyse des Mikrozensus nur vor dem Hintergrund der besonderen rechtlichen Reglementierung der amtlichen Statistik in Deutschland bewertet werden. Erfreulicherweise ist es aber aufgrund neuer rechtlicher Regeln möglich geworden, die spezifischen Variablen der EU-Arbeitskräfteerhebung ab dem Mikrozensus 1996 an die Forschung weiterzugeben. Es ist zur Zeit noch offen, wann dieser aktuelle Datensatz unter ähnlichen Nutzungsbedingungen wie der Mikrozensus 1995 für die Forschung zur Verfügung steht.

Für die Diskussion über Analysemöglichkeiten und methodische Fragen des Mikrozensus und als Forum für den Erfahrungsaustausch zwischen den Interessenten und Nutzern des Mikrozensus und der amtlichen Statistik kann die Konferenz "Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und Arbeitsmarkt" (siehe Ankündigung in diesem Heft) dienen.

Anmerkungen

1) Zudem sind die mit dem großen Stichprobenumfang verbundenen geringen Schätzfehler für die Messung von Veränderungstrends von großer Bedeutung. Auf die Möglichkeiten, die sich durch die Zusammenführung von Erhebungen mehrerer Zeitpunkte bieten, wird in diesem Aufsatz aus Platzgründen nicht eingegangen. Die Entwicklung des Mikrozensus seit 1957 ist in Emmerling/Riede 1997 dargestellt und die

Auswertungsmöglichkeiten von Daten aus 40 Jahren Mikrozensus sind in einem kürzlich vom Statistischen Bundesamt (1998) herausgegebenen Band dokumentiert.

- 2) Der Erwerbstyp wird in Kombination mit dem Unterhaltstyp in Variable EF34 wiedergegeben. Die für die Abgrenzung des Erwerbstyps relevanten Variablen sind EF61, EF63, EF65, EF66 und EF139. Angaben zur Verfügbarkeit für eine Arbeitsstelle sind in EF68 enthalten.
- 3) Die Abgrenzung der ILO-Erwerbslosen erfolgte in Anlehnung an die Methode der EU-Arbeitskräfteerhebung (vgl. Eurostat 1996). Veröffentlichte Ergebnisse von Eurostat beziehen sich i.d.R. auf die Bevölkerung in Privathaushalten ohne Wehrpflichtige am Familienwohnsitz. In den anonymisierten Einzeldaten sind Wehrpflichtige mit Angestellten zusammengefaßt, so daß eine zu Eurostat vergleichbare Auswertung nur eingeschränkt möglich ist.
- 4) Ab 1996 wird mit der neuen, vierten Leitfrage explizit nach ruhenden Beschäftigungsverhältnissen (z.B. bei Personen im Erziehungsurlaub) gefragt.
- 5) In Einkommensanalysen bzw. Regressionen mit gruppierten Daten wird in der Regel von den Klassenmitten ausgegangen und die offene Randklasse auf das Eineinhalbfache der Untergrenze gesetzt. Das Programmpaket Limdep 7.0 bietet mit „Grouped Data Regression Model“ eine eigenständige Schätzmethode für diese Art von Daten. Häufig unterscheiden sich jedoch die Koeffizienten aus OLS-Regressionen mit gruppierten und metrischen Einkommensangaben als abhängiger Variable nur geringfügig (Armingier 1995: 165).
- 6) Genauere Aussagen sind nicht möglich, da offensichtlich im Mikrozensus häufig ein Mann als Bezugsperson des Haushalts eingetragen ist, auch wenn er nicht die Antworten gegeben hat (vgl. Statistisches Bundesamt 1991). Im österreichischen Mikrozensus und im britischen Labour Force Survey liegt der Anteil an Selbstauskünften bei ungefähr zwei Drittel.
- 7) Mit dem in Abbildung 3 dargestellten Ergebnis wäre eventuell auch eine andere Erklärung kompatibel, nämlich daß Frauen das Einkommen der Männer richtig einschätzen, jedoch deren Status überschätzen.
- 8) An dieser Stelle ist auch darauf hinzuweisen, daß Umsiedler und Spätaussiedler bis auf wenige Ausnahmen als Deutsche gelten, so daß diese Gruppen im Mikrozensus nicht identifiziert werden können.
- 9) Diese Unterscheidung basiert auf der dreistelligen Berufsvariablen - kodiert nach der Klassifizierung der Berufe, Juli 1992 - und erfolgt in Anlehnung an eine Berufsklassifikation von Blossfeld (1983: 208). Beispielsweise werden Bauhilfsarbeiter der Kategorie einfacher manueller Berufe zugeordnet, während Maurer zu den qualifizierten manuellen Berufen gehören.
- 10) Ausnahmen sind erhebungstechnische Variablen, die jedoch nicht im faktisch anonymisierten Datensatz enthalten sind: Zahl der seit letzter Erhebung fortgezogenen

und verstorbenen Haushaltsmitglieder; Zahl der Haushalte in der Wohnung; Baualter der Wohnung (nur bei erstmaliger Befragung). Erstmals wieder seit 1982 wird ab 1996 das Haushaltseinkommen direkt erfragt.

11) Für die Ermittlung von Nichtehelichen Lebensgemeinschaften wurde bis einschließlich 1995 eine Typisierung verwendet. Ein Vergleich mit Ergebnissen aus der im Mikrozensus 1996 direkt gestellten Frage nach einer Lebenspartnerschaft zeigte eine weitgehende Übereinstimmung mit dem früheren Schätzverfahren. Ich danke Stefan Gruber (Statistisches Bundesamt), der seine Ergebnisse freundlicherweise zur Verfügung gestellt hat. An dieser Stelle danke ich auch Sabine Allafi und Hans-Joachim Heidenreich vom Statistischen Bundesamt für weitere hilfreiche Informationen zum Mikrozensus.

12) Während von 1982 bis 1995 diese Frage nach der Stellung im Betrieb im zweijährigen Turnus gestellt wurde, erfolgt nach dem neuen Mikrozensusgesetz die Erhebung dieses Merkmals nur noch im Abstand von vier Jahren. Diese aus Kostengründen vorgenommene Kürzung hat zur Folge, daß entsprechende Angaben nach 1996 erst wieder im Jahre 2000 vorliegen.

13) Die Vergleichbarkeit der Klassifizierung der Berufe (Ausgabe 1992) mit der Internationalen Standardklassifikation der Berufe (ISCO-88) ist nur auf der Basis des Vierstellers der Klassifizierung der Berufe möglich (Macht 1992). Im anonymisierten Grundfile liegen die Berufsangaben jedoch nur als Dreisteller vor, so daß es nicht möglich ist, diese Variable auf ISCO-88 bzw. ISCO-88 (COM) umzusetzen. ISCO-88 (COM) entspricht einer leicht modifizierten Fassung von ISCO-88, die insbesondere für die Anwendung auf die Arbeitskräfteerhebungen der EU zugeschnitten ist (Elias/Birch 1994).

14) Bei der Ziehung der anonymisierten 70%-Substichprobe aus dem Originalmaterial wurden aus diesem Grund Schichtungsfaktoren berücksichtigt. Die zufallsgesteuerte Auswahl der Haushalte erfolgte mittels Anordnung der Sätze nach regionalen Merkmalen (Bundesland, Regierungsbezirk und Gemeindegröße) und der Haushaltsgröße. Von einem für die Substichprobe gezogenen Haushalt wurden alle Personen in die Stichprobe übernommen. Für Personen in Gemeinschaftsunterkünften wurde jeweils eine eigene („Haushalts-“) Nummer vergeben.

15) Tatsächlich handelt es sich um ein zweistufiges Verfahren, bei dem in der ersten Stufe vorhandene Informationen über die ausgefallenen Haushalte zur Bildung von Haushaltsgewichten (Kompensation) benutzt werden. Die auf Personenebene gebildeten Soll-durch-Ist Gewichte der zweiten Stufe werden mit dem Kompensationsfaktor multipliziert und ergeben den endgültigen Hochrechnungsfaktor. Die Hochrechnungsklassen ergeben sich als Kombination von Geschlecht und Staatsangehörigkeit (Deutsche/Ausländer), ergänzt um Soldaten und Wehrpflichtige. Ausführlich informiert Heidenreich (1994) zur Hochrechnung des Mikrozensus.

16) Dieses Gewicht entspricht dem arithmetischen Mittel der Personengewichte im Haushalt. Seine Verwendung erbringt im wesentlichen die gleichen Ergebnisse, die mit dem Personengewicht erzielt werden, das aber aus Anonymisierungsgründen nicht in das Grundfile übernommen wurde.

17) Als Hinweis auf solche Probleme von Melderegistern kann evtl. ein Befund bei der Feldarbeit zum ALLBUS 1994 (Westdeutschland), der auf Basis einer Melderegisterstichprobe durchgeführt wurde, gewertet werden: Von den Interviewern wurden für Adressen von Ausländern drei- bis viermal häufiger als für Adressen von Deutschen angegeben, die Adresse sei falsch oder die Zielperson wäre verzogen. Für diese Informationen danke ich Achim Koch (ZUMA).

Literatur

- Alba, R./Müller, W./Schimpl-Neimanns, B., 1994a: Secondary Analysis of Official Microdata. S. 57-78 in: I. Borg/P.Ph. Mohler (Hrsg.), Trends and Perspectives in Empirical Social Research. Berlin: de Gruyter.
- Alba, R./Handl, J./Müller, W., 1994b: Ethnische Ungleichheit im deutschen Bildungssystem. Kölner Zeitschrift für Soziologie und Sozialpsychologie 46(2), 209-237.
- Arminger, G., 1995: Specification and Estimation of Mean Structures: Regression Models. S. 77-183 in: G. Arminger/C.C. Clogg/M.E. Sobel (Hrsg.), Handbook of Statistical Modeling for the Social and Behavioral Sciences. New York: Plenum Press.
- Berié, H./Mayer, H.-L., 1983: Entschließungen. Bundesarbeitsblatt 7-8: 15-22.
- Berk, R.A., 1983: An introduction to sample selection bias in sociological data. American Sociological Review 48: 386-398.
- Blair, J./Menon, G./Bickart, B., 1991: Measurement Effects in Self vs. Proxy Responses to Survey Questions: An Information-Processing Perspective. S. 145-166 in: P.P. Biemer/R.M. Groves/L.E. Lyberg/N.A. Mathiowetz/S. Sudman (Hrsg.), Measurement Errors in Surveys. New York: Wiley.
- Blossfeld, H.-P., 1983: Höherqualifizierung und Verdrängung - Konsequenzen der Bildungsexpansion in den siebziger Jahren. S. 184-240 in: M. Haller/W. Müller (Hrsg.), Beschäftigungssystem im gesellschaftlichen Wandel. Frankfurt: Campus.
- Dawe, F./Knight, I., 1997: A study of proxy response on the Labour Force Survey. Survey Methodology Bulletin 40(1): 30-36.
- DIW 1997: Erwerbsstatistik unterschätzt Beschäftigung um 2 Millionen Personen. DIW-Wochenbericht 64(38): 689-694.
- Elias, P./Birch, M., 1994: Errichtung einer EG-weiten Statistik der Berufe. ISCO 88 (COM). Fassung der Internationalen Standardklassifikation der Berufe 1988 zur Verwendung innerhalb der Europäischen Gemeinschaft. University of Warwick: Institute of Employment Research. [WWW-Dokument: <http://www.warwick.ac.uk/ier>]
- Emmerling, D./Riede, T., 1994: Zur Freiwilligkeit in der Auskunftserteilung im Mikrozensus. Wirtschaft und Statistik (6): 435-449.
- Emmerling, D./Riede, T., 1997: 40 Jahre Mikrozensus. Wirtschaft und Statistik (3): 160-174.
- Esser, H./Grohmann, H./Müller, W./Schäffer, K.-A., 1989: Mikrozensus im Wandel. Untersuchungen und Empfehlungen zur inhaltlichen und methodischen Gestaltung. Bd. 11 der Schriftenreihe „Forum der Bundesstatistik“. Hrsg. vom Statistischen Bundesamt. Stuttgart: Metzler-Poeschel.

- Eurostat 1996: Europäische Erhebung über Arbeitskräfte. Methodik und Definitionen. Themenkreis 3, Reihe E. Luxemburg.
- Gruber, S., 1995: Die Arbeitskräfteerhebung der Europäischen Union. *Wirtschaft und Statistik* (7): 518-529.
- Handl, J., 1996: Hat sich die berufliche Wertigkeit der Bildungsabschlüsse in den achtziger Jahren verringert? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 48(2): 249-273.
- Heidenreich, H.-J., 1994: Hochrechnung des Mikrozensus ab 1990. S. 112-123 in: S. Gabler/J. Hoffmeyer-Zlotnik/D. Krebs (Hrsg.), *Gewichtung in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Höhn, C., 1988: Zur Bedeutung von Definitionen und Erhebungskonzepten für demographische Analysen. S. 55-68 in: C. Höhn (Hrsg.), *Demographie in der Bundesrepublik Deutschland*. Bd. 18 der Schriftenreihe des Bundesinstituts für Bevölkerungsforschung. Boppard am Rhein: Boldt.
- Kohler, H./Rudolph, H./Spitznagel, E., 1996: Umfang, Struktur und Entwicklung der geringfügigen Beschäftigung - Eine Bestandsaufnahme. IAB-Kurzbericht 2/31.1.1996. Nürnberg.
- Krug, W./Nourney, M./Schmidt, J., 1994: *Wirtschafts- und Sozialstatistik. Gewinnung von Daten*. München: Oldenbourg (3. völlig neu bearb. Auflage).
- Lee, E.S./Forthofer, R.N./Lorimor, R.J., 1989: *Analyzing Complex Survey Data*. Series Quant. Applic. Soc. Sci.; 71. Newbury Park: Sage
- Little, R.J.A./Schenker, N., 1995: Missing Data. S. 39-77 in: G. Arminger/C.C. Clogg/M.E. Sobel (Hrsg.), *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum Press.
- Lüttinger, P./Riede, T., 1997: Der Mikrozensus. Amtliche Daten für die Sozialforschung. *ZUMA-Nachrichten* 41: 19-43.
- Macht, A. 1992: Klassifizierung der Berufe 1992. *Wirtschaft und Statistik* (12): 855-863.
- Meyer, K., 1994: Zum Auswahlplan des Mikrozensus ab 1990. S. 106-111 in: S. Gabler/J. Hoffmeyer-Zlotnik/D. Krebs (Hrsg.), *Gewichtung in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Recktenwald, J., 1997: International vergleichende Messung von Beschäftigung und Arbeitslosigkeit - Bestrebungen zur Harmonisierung in der Europäischen Union (EU). S. 77-86 in: Statistisches Bundesamt (Hrsg.), *Statistische Informationen zum Arbeitsmarkt - Konzepte und Kritik, Anwendung und Auslegung*. Bd. 30 der Schriftenreihe „Forum der Bundesstatistik“. Stuttgart: Metzler-Poeschel.

- Rendtel, U., 1991: Die Schätzung von Populationswerten in Panelerhebungen. *Allgemeines Statistisches Archiv* 75: 225-244.
- Riede, T./Emmerling, D., 1994: Analysen zur Freiwilligkeit der Auskunftserteilung im Mikrozensus. Sind Stichprobenergebnisse bei freiwilliger Auskunftserteilung verzerrt? *Wirtschaft und Statistik* (9): 733-742.
- Rinne, H., 1994: *Wirtschafts- und Bevölkerungsstatistik*. München: Oldenbourg.
- Rohwer, G. 1994: TDA Working Paper 1-3. Replication Methods for Variance Estimation. Universität Bremen.
- Schimpl-Neimanns, B./Lüttinger, P., 1993: Die Entwicklung bildungsspezifischer Ungleichheit: Bildungsforschung mit Daten der amtlichen Statistik. *ZUMA-Nachrichten* 32: 76-115.
- Statistisches Bundesamt, 1991: Zur Antwortbereitschaft von Haushalten am Beispiel der Mikrozensus-Testerhebung 1986. Kontrastgruppenanalyse und Logit-Modelle. Heft 11 der Schriftenreihe "Ausgewählte Arbeitsunterlagen zur Bundesstatistik". Wiesbaden.
- Statistisches Bundesamt, 1996: *Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihe 4.1.1 Stand und Entwicklung der Erwerbstätigkeit 1995 (Ergebnisse des Mikrozensus)*. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt, 1997: *Stellungnahme für die öffentliche Anhörung der Ausschüsse Arbeit und Sozialordnung sowie Familie, Senioren, Frauen und Jugend des Deutschen Bundestages am 10. Dezember 1997 zum Themenbereich I: Umfang, Struktur und Fehlentwicklung der sozialversicherungsfreien geringfügigen Beschäftigung*. Wiesbaden.
- Statistisches Bundesamt, 1998: *Im Blickpunkt: Leben und Arbeiten in Deutschland. 40 Jahre Mikrozensus*. Stuttgart: Metzler-Poeschel.
- Winship, C./Mare, R.D., 1989: Loglinear Models with Missing Data: A Latent Class Approach. S. 313-368 in: C. Clogg (Hrsg.), *Sociological Methodology*; 19. Oxford: Basil Blackwell.
- Wirth, H./Lüttinger, P., 1998: Klassenspezifische Heiratsbeziehungen im Wandel? Die Klassenzugehörigkeit von Ehepartnern 1970 und 1993. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Heft 1: 47-77.
- Wolter, K.M., 1985: *Introduction to Variance Estimation*. Heidelberg: Springer.

„WIENERWALD AUF MALLORCA“: NEGATIVE EMOTIONEN VON DEUTSCHEN GEGENÜBER IHRER EIGENEN NATION

THOMAS BLANK UND HORST-ALFRED HEINRICH

Die Vielfalt emotionaler Aspekte nationaler Identität werden in Untersuchungen zum Nationalstolz selten erschöpfend diskutiert. So wurden negative Emotionen zwar in theoretischen Überlegungen thematisiert, empirisch bisher aber in Ermangelung geeigneter Instrumente nicht systematisch untersucht. Wir stellen hier die Entwicklung eines Meßinstrumentes vor, welches als ein erster Ansatzpunkt für derartige systematische Untersuchungen dienen könnte. Dazu wurde zunächst mittels einer offenen Frage das Peinlichkeitsgefühl gegenüber der eigenen Nation erhoben. Aus den Antworten wurde dann ein geschlossenes Meßinstrument zur Erfassung dieser Emotion entwickelt. Zur externen Validierung des Instrumentes wurden zwei Telefonbefragungen durchgeführt. Die Analysen ergeben ein vierfaktorielles Peinlichkeitsmodell. Die externe Validierung zeigt, daß negative Emotionen gegenüber der eigenen Nation als ein eigenständiger Aspekt nationaler Identität zu betrachten sind.

Investigations of national identity do not usually deal with other emotional aspects than national pride. The paper reports on the development of a closed format instrument based on answers to an open question on feelings of embarrassment with respect to one's nation. Two telephone surveys served to validate the instrument externally. Our findings point to a four-factor measurement model of embarrassment. External validation shows that negative emotions with respect to one's own nation are a specific aspect of national identity.

1. Einleitung und Problemstellung¹⁾

Die Frage nach dem Zusammenhalt der Deutschen sowie nach der Bedeutung nationaler Werte, Symbole und Normen nimmt seit Mitte der 80er Jahre einen zentralen Stellenwert sowohl im politischen als auch im gesellschaftswissenschaftlichen Diskurs ein (vgl. z.B. den Literaturüberblick bei Zitelmann 1990; Giesen/Leggewie 1991). Mit

der friedlichen Revolution in der DDR und der Vereinigung beider deutscher Staaten erhielten diese Diskussionen weiteren Auftrieb. In ihnen geht es einerseits um die internationale Rolle Deutschlands. Hier stellt sich die Frage, ob sich das künftige deutsche Selbstverständnis auf außenpolitische Machtdemonstration oder Fortsetzung der europäischen Integration hin ausrichten soll. Andererseits zielt die Debatte auf das Binnenverhältnis der Gesellschaftsmitglieder zueinander, ein Punkt der im Zusammenhang mit den Anschlägen gegen Ausländer aber auch mit deren rechtlichen Status zu sehen ist (vgl. Baring 1992; Bubis/Schäuble 1996).

In der Regel mangelt es den Publikationen zu dieser Thematik an einer trennscharf definierten Begrifflichkeit. Das gilt sowohl für den Terminus der nationalen Identität als auch für Begriffe wie Nationalbewußtsein, Nationalismus, Patriotismus und Nationalstolz, die nicht selten als synonym behandelt werden (vgl. Mummendey 1992; Noelle-Neumann/Köcher 1987; Topf et al. 1990; Herdegen 1987; Mohler/Götze 1992). Außerdem bleibt deren Bezug zur nationalen Identität ungeklärt. Auf diese Weise werden nicht nur unterschiedliche Sinngehalte miteinander vermischt, es werden auch so grundlegende Kategorien wie Kognition und Emotion konfundiert (Blank 1997). Wir wenden uns der Betrachtung des letztgenannten Aspekts zu und fragen nach der Verknüpfung zwischen Emotionen gegenüber dem nationalen Kollektiv und nationaler Identität.

Soweit emotionale Aspekte in der Identitätsdebatte bisher berücksichtigt wurden, bezogen sie sich auf die positive Bewertung des Kollektivs,²⁾ nämlich die Emotion des Stolzes (vgl. z.B. Noelle-Neumann/Köcher 1987; Mohler/Götze 1992). Beim Nationalstolz handelt es sich allerdings insoweit um ein problematisches Konstrukt, als positive Zusammenhänge mit der Überbewertung der eigenen Nation, mit Fremdenfeindlichkeit, mit Antisemitismus und mit einer Verdrängung und beziehungsweise oder Idealisierung der deutschen Geschichte bestehen (vgl. Blank/Schmidt 1993, 1997; Blank/Wittenberg/ Schneider 1998). In der politischen Theorie wird insbesondere die Frage der Beziehung zwischen Nationalstolz und verschiedenen Formen der Systemunterstützung und -stabilität diskutiert (Almond/Verba 1963; Topf et al. 1990; Westle 1994; Easton 1975).

Als Konzept, das ausdrücklich positive wie auch negative Einstellungen gegenüber dem Kollektiv zuläßt, ist der konstruktive Patriotismus (Staub 1997; Blank/Schmidt 1997; Blank/Wittenberg/Schneider 1998) zu betrachten. Im Unterschied zum blinden Patriotismus (Adorno et al. 1950) oder zum Nationalismus (Kosterman/Feshbach 1989) definiert Staub konstruktiven Patriotismus dadurch, daß dem Kollektiv die Gefolgschaft dann verweigert werde, wenn dessen Ziele destruktiv seien. Allerdings werden auch bei

diesem Konzept negative Emotionen gegenüber der Eigengruppe weder konzeptualisiert noch operationalisiert. Das gilt zum Beispiel für Gefühle wie: die Nation zu verachten, Angst vor der Nation zu haben, Abscheu gegenüber der Nation zu empfinden, sich für die Nation beziehungsweise als Mitglied eben dieser zu schämen oder das Gefühl von Peinlichkeit aufgrund kollektiver Ereignisse (vgl. Tajfel 1982).

Die genannten negativen Emotionen sind deshalb für nationale Identität relevant, weil sie in Verbindung mit positiven Empfindungen individuelle Ambivalenzen gegenüber dem Kollektiv zum Ausdruck bringen (Tajfel/Turner 1986). Das ausschließliche Beharren auf dem Nationalstolz suggeriert hingegen, ein Kollektiv vermittele allein positive, das Selbstvertrauen stärkende Gefühle (vgl. Noelle-Neumann/Köcher 1987). Ein solcher Ansatz blendet jedoch zentrale Erkenntnisse der Emotions- und Sozialpsychologie aus. Ihnen zufolge nehmen Individuen zeitgleich nicht nur verschiedene Emotionen mit unterschiedlicher Intensität wahr (Rahn et al. 1994). Vielmehr treten auch solche Gefühle parallel auf, die sich zwar widersprechen, sich aber ebenso gegenseitig bedingen beziehungsweise kontrollieren (vgl. Kruse 1985). Derartige Ambivalenzen sollten sich dann auch auf die nationale Identität und damit auf das Verhalten des Einzelnen gegenüber der Nation auswirken. Wir konzentrieren uns in diesem Beitrag auf negative emotionale Aspekte nationaler Identität, um über die bisherige Einseitigkeit in Form ausschließlich erfolgter Thematisierung von Nationalstolz hinauszukommen.

2. Theoretische Implikationen negativer Emotionen gegenüber der Eigengruppe

Psychologischen Konzepten zufolge setzen sich individuelle Einstellungen aus Kognitionen, Emotionen und Evaluationen gegenüber einem bestimmten Objekt zusammen (vgl. Rosenberg/Hovland 1960; Eagly/Chaiken 1993). Allen drei Aspekten kommt bei der Bildung von Einstellungen jeweils eine eigenständige Bedeutung zu (vgl. Ottati/Wyer 1993). Sie stehen jedoch auch in Wechselbeziehungen. So sind zum Beispiel Kognitionen auch emotional repräsentiert (vgl. Rosenberg/Hovland 1960) und Emotionen können über Kognitionen vermittelt sein (vgl. Vester 1991).

Die vorangestellten Grundannahmen lassen sich auf die nationale Identität des Individuums im Sinne einer allgemeinen Einstellung gegenüber der eigenen Nation übertragen (vgl. Zanna/Rempel 1988). So bestehen die kognitiven Aspekte zum Beispiel aus dem Wissen über die Gruppenzugehörigkeit, die Kriterien der Gruppenzugehörigkeit und die Gruppenspezifität (vgl. Schweigler 1973; Tajfel 1982; Wakenhut 1994; Blank 1997).

Als exemplarisch für negative Emotionen gegenüber der Eigengruppe kann die Wahrnehmung eines spezifischen Verhaltens deutscher Touristen im Ausland gelten. So empfinden es manche Deutschen als peinlich, wenn sich Nationsangehörige am Urlaubsort im Ausland ausfallend benehmen und sei es auch nur, wenn diese dieselben Verhältnisse wie in Deutschland, den „Wienerwald“ mit Schnitzel und Bier, einklagen (vgl. Skala 1994).³⁾ In einem solchen Fall offenbart sich für das Peinlichkeit empfindende Individuum ein Zwiespalt zwischen der Zugehörigkeit zur Eigengruppe einerseits und der Wahrnehmung des als anstößig empfundenen Verhaltens von Angehörigen der eigenen Nation andererseits. In Abhängigkeit von der konkreten Situation kann das dazu führen, daß sich die einzelne Person entscheidet, wie weit sie sich mit dem Kollektiv der Deutschen identifiziert, ihm zugehören möchte. Damit kann eine derartige negative Emotion ein Indikator für Distanzierung zur Gruppe und ihren Mitgliedern sein.

Gefühle gegenüber der Eigengruppe können sich auf verschiedene Aspekte der Nation beziehen. Gemeint sind damit auf nationaler Ebene Ereignisse, welche dieses Kollektiv betreffen, oder Verhaltensweisen von Repräsentanten der Eigengruppe, die als Objekte von Emotionen fungieren.

Darüber hinaus spielen die spezifischen Formen individueller und kollektiver Kognition und Evaluation gegenüber diesen Objekten eine Rolle. Das durch diese Größen beeinflusste subjektive Gefühl läßt sich auf einem emotionalen Kontinuum anordnen, das sich zwischen Nähe zum (positiv) und Distanzierung vom (negativ) Kollektiv bewegt.

Werden die Nation oder ihre Teilaspekte (Kollektivgüter; vgl. Olson 1968) als negativ empfunden, so geht dieses Gefühl mit entsprechenden Kognitionen einher. Zusammen mit diesen Kognitionen kann eine solche Emotion zudem Verhaltensrelevanz besitzen (Ajzen 1988). So kann das oben beschriebene Verhalten von Touristen negativ auf die nationale Identität wirken, indem es im Einzelnen den Wunsch zur Distanzierung vom Kollektiv auslöst. Auf der anderen Seite können negative Emotionen in bezug zu Kritik- und Veränderungspotential stehen (Hirschman 1970). Es ist denkbar, daß ein Individuum aufgrund negativer Erfahrungen den Anstoß erhält, in Zukunft bei einem ähnlichen Ereignis einzugreifen und dieses zu verhindern (Merton 1949; Lüdemann 1995).⁴⁾

Unabhängig von der sozialpsychologischen Debatte um das Verhältnis zwischen Emotion und Evaluation (vgl. Fishbein/Ajzen 1975; Fiske/Pavelchak 1986) sind hier für uns zwei Aspekte entscheidend:

1) Die Wahrnehmung spezifischer negativer Emotionen gegenüber der Nation (z.B. Angst, Furcht, Haß, Scham) steht in einem wechselseitigen Zusammenhang mit dem persönlichen Verhältnis zur Eigengruppe.

2) Dementsprechend steht die Stärke nationaler Identifikation in zweifachem Zusammenhang mit negativen Gefühlen gegenüber der Nation. Zum einen kann die Identifikation mit der Nation zunehmend schwächer ausfallen, je stärker der Wunsch nach Distanzierung vom Kollektiv aufgrund der mit ihm verbundenen negativen Emotionen wächst. Zum anderen ist aber eine identitätsstärkende Wirkung zu vermuten, die in der Motivation, einen Beitrag zur Veränderung der Ursachen negativer Emotionen zu leisten, begründet ist (vgl. Tajfel 1982; Hirschman 1970).

Derartige Differenzierungen sind bei positiven Emotionen (Stolz, Liebe) gegenüber der Eigengruppe nicht sinnvoll, da sie positiv mit nationaler Identität verknüpft sind (Blank/Schmidt 1997). So haben Topf et al. (1990) zeigen können, daß Personen ohne Stolz auf kollektive Güter ein ausgeprägteres Veränderungsbedürfnis im Hinblick auf das politische System haben. Sie stellen in Deutschland auch die größte Nichtwählergruppe im Hinblick auf die Unterscheidung zwischen Stolzen und Nichtstolzen.

Für die Konzeptionalisierung negativer Emotionen beschränken wir uns hier auf die Peinlichkeit, weil es sich unserer Ansicht nach um ein empirisch leicht zu erfassendes Gefühl handelt. Zudem wird die Äußerung einer solchen Emotion vermutlich sozial eher nicht sanktioniert (vgl. Reinecke 1991). Um dieses Gefühl näher zu analysieren, unterscheiden wir bei der folgenden generalisierten Konzeption drei Typen von Akteuren.

- A1: Akteur, dessen Verhalten hinsichtlich Peinlichkeit bewertet wird;
- A2: der Peinlichkeit empfindende Akteur;
- A3: die Referenznorm oder -gruppe, die von A2 als maßgebend angesehen wird.

Für die Beziehungsstruktur zwischen den Akteuren gelten folgende Regeln:

- A2 nimmt das Verhalten von A1 wahr;
- A3 ist für A2 als soziale Kontrollinstanz bedeutsam;
- A2 glaubt, daß A3 das Verhalten von A1 mißbilligt beziehungsweise daß dessen Verhalten gegen die Gruppennorm verstößt;
- A2 weiß sich als Mitglied derselben Gruppe wie A1;
- A2 weiß, daß es derselben Gruppe zugerechnet wird wie A1.

Sind diese Bedingungen erfüllt, empfindet A2 Peinlichkeit über das Verhalten von A1. Diese Struktur soll auch gültig sein, wenn eine Normverletzung durch die Gruppe wahr-

genommen wird. Das Peinlichkeitsgefühl ist damit indirekt auch Ausdruck eines kollektiven Selbstverständnisses. Auf diese Ebene übertragen glauben Individuen - soweit sie eine verletzte Norm internalisiert haben -, daß ihnen durch Dritte Verantwortung für wahrgenommene Normüberschreitungen zugeschrieben wird, obwohl sie diese Verantwortung für sich selbst nicht erleben.

Ausgehend von diesem Peinlichkeitskonzept skizzieren wir nachfolgend unsere empirische Umsetzung und die Entwicklung des Meßinstruments (Abschnitt 3). In Abschnitt 4 wird die interne Validität unserer Peinlichkeitsskala geprüft. Die Mehrdimensionalität des Konstrukts gibt einerseits Hinweise auf unterschiedliche Akte gesellschaftlichen Handelns, welche zu einem Peinlichkeitsgefühl führen. Andererseits zeigt sich ein Zusammenhang zwischen diesen Handlungsakten und der Peinlichkeit der Gruppenmitgliedschaft. Schließlich ist auch eine subgruppenspezifische Peinlichkeitsdimension nachweisbar.

Im Rahmen der externen Validierung wird im fünften Abschnitt der Zusammenhang zwischen dem Peinlichkeitsgefühl und Einstellungen gegenüber Fremdgruppen und zur Salienz der Gruppenzugehörigkeit geprüft. Anhand dieser Kriterien wird theoriegeleitet die inhaltliche Gültigkeit der Peinlichkeitsdimensionen geprüft. Im letzten Abschnitt diskutieren wir dann die Ergebnisse.

3. Stichproben und Instrumentenentwicklung

Zur Entwicklung und Validierung unserer Operationalisierung wurden drei unabhängige repräsentative regionale Stichproben verwendet. Die erste wurde in Gießen im April 1992 (GI, N = 182), die zweite in Münster im Herbst 1993 (MS-I, N = 300) und die dritte ebendort im Frühjahr 1994 (MS-II, N = 308) erhoben (vgl. Blank et al. 1993; Blank 1994; Blank et al. 1994).

Zur Messung des auf das Kollektiv bezogenen Peinlichkeitsgefühls lag keine standardisierte Skala vor. Deshalb wurde in Gießen eine offene Frage nach dem Peinlichkeitsempfinden gestellt, welches entsteht, wenn die nationale Zugehörigkeit als Identifikationsrahmen bedeutsam wird.⁵⁾ Für die systematische Auswertung der von 176 Personen geäußerten 278 unterschiedlichen Begriffe wurde ein Kategorienschema erstellt (vgl. Merten 1983: 302-308). Es schloß sich ein Ratingverfahren an, in dem die Rater alle genannten Begriffe möglichst einheitlich dem Kategorienschema zuordnen sollten.⁶⁾ Insgesamt ließen sich von den 278 unterschiedlichen Begriffen 77,3 Prozent eindeutig einer Kategorie zuordnen.⁷⁾ Im Ergebnis sahen die Gießener Befragten folgende Kategorien jeweils als unangenehm an (in Klammern sind jeweils Textbeispiele angegeben):

- 25 Prozent die derzeitige Ausländerfeindlichkeit
(„Ausländerhaß“; „Anschläge auf Asylantenheime“),
- 17 Prozent unangemessenes Verhalten deutscher Touristen im Ausland,
(„Deutsche im Ausland“; „Auftreten deutscher Touristen im Ausland“),

- 17 Prozent als typisch angesehene deutsche Charaktereigenschaften („Besserwisserei“; „Spießbürgertum“),
- 13 Prozent den derzeitigen Rechtsradikalismus („Neonazis“; „Aufkleber: Ich bin stolz Deutscher zu sein“),
- 13 Prozent die derzeitige Politik („Scheckbuchpolitik“; „Eskapaden von unserem tollen Bundeskanzler“) und
- 10 Prozent die nationalsozialistische Vergangenheit („Zweiter Weltkrieg“; „Umgang im Dritten Reich mit Juden“).

Damit zeigt sich, daß sich das hier zur Diskussion stehende negative Gefühl in hohem Maße einerseits auf die 1992 aktuellen gegen Fremdgruppen gerichteten Ereignisse und andererseits auf die Wahrnehmung von Teilaspekten der Eigengruppe bezieht. Letztere sind zu unterteilen in allgemeine Charakteristika sowie die Erinnerung an die kollektive Vergangenheit.

Dem Kategorienschema folgend wurde eine standardisierte Skala entsprechend dem Sinngehalt der einzelnen Kategorien entwickelt. Zusammen mit der Replikation der offenen Frage wurde sie im Politbarometer MS-I (vgl. Blank 1994) erhoben. Aufgrund der Ergebnisse der Kreuzvalidierung mußten einige Itemformulierungen überarbeitet werden. Die reformulierten Items wurden dann im Politbarometer MS-II (vgl. Blank et al. 1994) erneut erhoben (vgl. Tabelle 1).

Die Auswertung der dritten Umfrage zeigt auf der deskriptiven Ebene, daß, wie aus den Mittelwerten in Tabelle 1 ersichtlich ist, den Befragten in erster Linie drei gegenwarts-, eine vergangenheitsbezogene und eine allgemeine Kategorie peinlich sind. Als sehr unangenehm sehen sie den 1994 aktuellen Rechtsradikalismus, gewalttätige Ausschreitungen gegenüber Ausländern und die aktuelle Ausländerfeindlichkeit sowie die Judenverfolgung im Dritten Reich an. Ähnlich peinlich ist ihnen das Verhalten deutscher Touristen im Ausland. Darüber hinaus ist in abgeschwächter Form auch die deutsche Geschichte mit einem negativen Gefühl behaftet. Eher nicht peinlich ist dagegen den Befragten die Thematisierung der Zugehörigkeit zur Gruppe der Deutschen. Das Resultat gibt somit einen Hinweis darauf, daß die Identifikation mit der nationalen Eigengruppe mit ambivalenten Bewertungen gruppenspezifischer Teilaspekte, nämlich der Fremdgruppenablehnung, im Sinne des konstruktiven Patriotismus (vgl. Staub 1997; Blank/Schmidt 1997) verbunden sein kann.

Weiterhin ist die Identifikation mit Deutschland bei den Befragten nicht zwangsläufig mit einer generalisierten vorbehaltlosen positiven Bewertung oder gar Überbewertung im Sinne des Nationalismus (vgl. Kosterman/Feshbach 1987; Blank/Schmidt 1997) verbunden. Aus dem Antwortverhalten läßt sich zunächst auf zwei latente Dimensionen

negativen Empfindens schließen: a) die Peinlichkeit der eigenen Gruppenzugehörigkeit und b) die der Fremdgruppenablehnung. Insofern zeigen sich damit die schon bei Levinson (1950) und Tajfel (1982) relevanten Teilaspekte sozialer Identität. Darüber hinaus kann auf einen weiteren Aspekt geschlossen werden: c) unangenehme Emotionen, die aus der Geschichte der Eigengruppe resultieren.

Tabelle 1: Die standardisierten Peinlichkeitsitems: Formulierungen, Mittelwerte und Standardabweichungen (MS-II, N = 308)

Skalierung: 1 = "überhaupt nicht peinlich"; 7 = "sehr peinlich"			
Label	Itemformulierung	x	s
dt.Tourist	Wie peinlich sind Ihnen persönlich deutsche Touristen im Ausland, die sich dort respektlos verhalten?	6.21	1.44
dt.Geschichte	Wie peinlich ist Ihnen persönlich die deutsche Geschichte?	3.48	1.99
dt.angesehen	Wie peinlich ist es Ihnen persönlich, wenn Sie darauf angesprochen werden, daß Sie Deutsche/r sind?	1.85	1.47
Ausl_Feindschft	Wie peinlich ist Ihnen persönlich die in Deutschland herrschende Ausländerfeindlichkeit?	6.14	1.55
Jud_Verfolgg	Wie peinlich ist Ihnen persönlich die Judenverfolgung in der Zeit des Nationalsozialismus?	5.96	1.86
dt.Herkunft	Wie peinlich ist es Ihnen persönlich, daß Sie aus Deutschland kommen?	1.65	1.18
Gewalt_gg.Ausl	Wie peinlich sind Ihnen persönlich die derzeitigen gewalttätigen Ausschreitungen gegen Ausländer in Deutschland?	6.47	1.21
Niederlg.1945	Wie peinlich ist Ihnen persönlich die Niederlage Deutschlands 1945?	1.76	1.59
Ordnngssinn	Wie peinlich ist Ihnen persönlich der Ordnungssinn der Deutschen?	2.90	1.83
Rechtsradikal	Wie peinlich ist Ihnen persönlich der derzeitige Rechtsradikalismus in Deutschland?	6.37	1.35
Dt.-Sein	Wie peinlich ist es Ihnen persönlich, daß Sie Deutsche/r sind?	1.61	1.19
Umgang_NS	Wie peinlich ist es Ihnen persönlich, daß noch immer	2.39	1.91

	viel über den Nationalsozialismus geredet wird?		
--	---	--	--

Die polychorischen Korrelationen⁸⁾ lassen die drei oben genannten latenten Konstrukte von Peinlichkeit erkennen. So korrelieren die drei Indikatoren, die das individuelle Verhältnis zur Eigengruppe thematisieren: (dt.angesehen, dt.Herkunft und Dt.-Sein) hoch positiv miteinander. Dieses konsistente Antwortverhalten zeugt von einer eindeutigen Einstellung der Befragten bezüglich der Gruppenzugehörigkeit. Dasselbe gilt für die Indikatoren zur Bewertung des Verhältnisses gegenüber Fremdgruppen (Ausl-Feindschaft, Gewalt_gg.Ausl, Rechtsradikal, Jud_Verfolg, dt.Tourist) mit ebenfalls hohen positiven Korrelationen. In beiden Fällen ist es zulässig, auf eine latente Dimension zu schließen.

Die Items, die das Verhältnis zur deutschen Geschichte thematisieren, weisen im Gegensatz dazu eine eher diffuse Struktur auf. Das Item zur deutschen Geschichte (dt.Geschichte) korreliert sowohl mit den Indikatoren der Eigen- wie auch mit denen der Fremdgruppenbewertung. Die höchste positive Korrelation zeigt sich zwischen den negativen Empfindungen zur deutschen Geschichte und zur Judenverfolgung, was plausibel und ein Indiz dafür ist, daß die Shoah bei der Geschichtsbewertung in Deutschland eine wichtige Rolle spielt. Die beiden anderen Indikatoren, zum einen das Wissen um die deutsche Kapitulation am Ende des Zweiten Weltkrieges (Niederlg.1945) sowie zum anderen zum Diskurs über den Nationalsozialismus (Umgang_NS) korrelieren intern miteinander, weisen mit allen anderen Variablen aber abweichende Muster auf. Umgang_NS steht mit den übrigen Items in keiner Beziehung. Niederlg.1945 korreliert mit den anderen durchgängig negativ, so auch mit dem Item zur deutschen Geschichte allgemein (-.193). Daher ist zu vermuten, daß ein vierter latenter Faktor vorliegt: Die Geschichtsbewertung durch politisch rechts Orientierte. Wem die Niederlage Deutschlands 1945 peinlich ist, der empfindet es auch als unangenehm, daß noch immer viel über den Nationalsozialismus geredet wird. Dieser Personenkreis wird deshalb als rechtsorientiert eingeschätzt, weil diejenigen, die dieser Meinung zuzurechnen sind, die deutsche Geschichte allgemein, die Ausländerfeindlichkeit sowie die Judenverfolgung in Deutschland nicht als unangenehm empfinden. Im folgenden wird zu zeigen sein, ob diese spezifische negative Empfindung vor dem Hintergrund einer Verleugnung von Teilen der deutschen Geschichte und ihren Konsequenzen (vgl. Giordano 1987) zu interpretieren ist.

4. Multidimensionalität der Peinlichkeitsbewertung

Um die interne Validität der Indikatoren zu prüfen, wurden die Items einer konfirmatorischen Faktorenanalyse unterzogen (LISREL 8, vgl. Jöreskog/Sörbom 1993a, 1993b). In Tabelle 2 ist die vollstandardisierte Lösung des akzeptierten Modells dargestellt. Die

standardisierten Faktorenladungen in diesem Modell werden von uns als standardisierte Validitätskoeffizienten interpretiert (vgl. Bollen 1989).⁹⁾

Tabelle 2: Vollstandardisierte Faktorenladungen der Peinlichkeitsindikatoren

	Latente Peinlichkeitsfaktoren				Meßfehler
	P-1	P-2	P-3	P-4	
Ordnungssinn	.24	.46			.73
Ausl_Feindschft	.72	.34			.37
dt.Tourist	.62				.62
Gewalt_gg.Ausl	.95				.09
Rechtsradikal	.89				.21
Jud_Verfolgg	.56		.35		.34
dt.angesehen		.87			.24
dt.Herkunft		.98			.04
Dt.-Sein		.97			.06
dt.Geschichte			.89		.21
Niederlg.1945			-.20	.42	.79
Umgang_NS				1.00	--

Chi-Quadrat = 39.84, df = 49, p = .82, GFI = .99.

Der Faktor P-1 wird in erster Linie durch die Indikatoren *Gewalt_gg.Ausl*, *Rechtsradikal*, *Auslfeindschft* und *dt.Tourist* indiziert (vgl. Tabelle 2). Folglich wird das Wissen um diskriminierendes Verhalten gegenüber Fremdgruppen als peinlich beziehungsweise als unangenehm empfunden. Diese Emotion bezieht sich hauptsächlich auf die in Deutschland mit Gewalt ausgeübten Handlungen. Aber auch die dahinterstehenden Fremdgruppen abwertenden Überzeugungen werden als peinlich empfunden. Umgekehrt ist bei den Befragten die Verletzung fremdländischer kultureller und religiöser Regeln sowie die Betonung deutschen Lebensstils im Ausland ebenfalls mit dem Gefühl von Peinlichkeit verknüpft. Das läßt den Schluß zu, derartiges Verhalten gegenüber den Gastkulturen wird als Diskriminierung eingestuft. Diese Folgerung wird bestätigt durch die hohen Mittelwerte der hier relevanten Indikatoren (vgl. Tabelle 1). Demnach ist bei den Befragten von einem kollektiven Konsens dieses Peinlichkeitsgefühls auszugehen.

Der Faktor P-2 besteht in erster Linie aus den drei Indikatoren *dt.angesehen*, *dt.Herkunft* und *Dt.-Sein*. Er mißt den Grad der in bezug auf die Gruppenzugehörigkeit

empfundene Peinlichkeit. Dieser Faktor kann damit als genereller Peinlichkeitsfaktor hinsichtlich der nationalen Identifikation betrachtet werden. Den Befragten ist es nach den Mittelwerten mehrheitlich nicht peinlich, wenn diese Gruppenzugehörigkeit thematisiert wird (vgl. Tabelle 1). Etwas schlechter laden auf diesem Faktor die Items *Ordnungssinn* und *Ausl_Feindschft*. Beide Indikatoren werden damit auch als weitere Aspekte des Eigengruppenimages thematisiert. Denjenigen, denen das Deutschsein nicht peinlich ist, ist auch die Fremdenfeindlichkeit und der Ordnungssinn nicht peinlich und umgekehrt. Insofern kann dieser Faktor nach Tajfel (1982) und Blank/Schmidt (1993, 1997) auch als ein emotionaler Aspekt von Nationalismus im Sinne von Kosterman/Feshbach (1989) interpretiert werden.

Der Faktor P-3 wird in erster Linie durch die Peinlichkeit der Eigengruppengeschichte *dt.Geschichte* bestimmt. Das heterogene Verhältnis der Befragten zur Geschichte Deutschlands zeigt sich empirisch daran, daß die beiden anderen Indikatoren *Jud_Verfolg* und *Niederlg.1945* diesen Faktor deutlich schlechter abbilden. Zudem thematisieren diese beiden Items ein jeweils unterschiedliches Geschichtsverständnis. Wem die deutsche Geschichte peinlich ist, der hat auch eher ein unangenehmes Empfinden bei der Erinnerung an die Verfolgung der Juden, in bezug auf die Niederlage 1945 dagegen eher nicht. Der Nationalsozialismus und die Shoah sind offensichtlich Bezugspunkte einer auf die Geschichte bezogenen negativen Emotion. Die Meßqualität der Indikatoren läßt aber offen, ob das Peinlichkeitsgefühl eine eher schlechte Beschreibung dieses Gefühls ist und welche Aspekte bei der Geschichtsbewertung im Mittelpunkt stehen.

Der Faktor P-4 thematisiert einen spezifischen Aspekt der Peinlichkeit deutscher Geschichte: die nicht aufgearbeitete Vergangenheit. Die Qualität der Indikatoren ist zwar nicht befriedigend, eine inhaltliche Interpretation aber trotzdem möglich. Wem die Niederlage 1945 peinlich ist, der empfindet es auch als unangenehm, daß der Nationalsozialismus nach wie vor thematisiert wird. Hier dürfte es sich um das spezifische Peinlichkeitsempfinden politisch Rechtsorientierter handeln.

Tabelle 3: Vollstandardisierte Faktorkorrelationen der Peinlichkeitsfaktoren

	P-1	P-2	P-3	P-4
P-1	1.00			
P-2	--	1.00		
P-3	.55	.54	1.00	
P-4	--	--	--	1.00

Betrachtet man die Faktorkorrelationen (vgl. Tabelle 3), so zeigen sich nur zwei signifikante Beziehungen. Die Peinlichkeit der deutschen Geschichte insgesamt korreliert sowohl mit dem unangenehmen Gefühl hinsichtlich der Eigengruppenbewertung wie auch mit demjenigen des Verhaltens gegenüber Fremdgruppen positiv. Auf der latenten Ebene bedeutet dies: Wem die deutsche Geschichte insgesamt peinlich ist, dem ist auch die Gruppenzugehörigkeit sowie die Diskriminierung von Fremdgruppen unangenehm. Eine kritische emotionale Bewertung der deutschen Geschichte spiegelt sich danach in Identifikationsproblemen mit der Eigengruppe wider. Negative Emotionen gegenüber der Eigengruppe haben somit eine Bedeutung für die Identifikation mit eben dieser.

Die Nullkorrelationen zwischen den Faktoren sind bei zwei der vier ebenfalls inhaltlich interpretierbar:

a) Die emotionale negative Bewertung (P-1) der Fremdgruppenabwertung wird vor dem Hintergrund allgemeiner Werthaltungen vorgenommen, die nicht spezifisch mit deutscher Identität (P-2) verbunden sind. Die Gruppenzugehörigkeit hat hier keine Bedeutung. Die historische Erfahrung mit der Shoah wird dagegen deutlich als belastender Teil deutscher Identität begriffen.¹⁰⁾

b) Die rechtsorientierte spezifische Vergangenheitsbewertung (P-4) kann im Sinne einer Verleugnung der nationalsozialistischen Vergangenheit interpretiert werden. Bislang wurde diese These (vgl. Giordano 1987) über den Umgang mit diesem Teil der deutschen Geschichte für positive Emotionen bestätigt (vgl. Blank/Schmidt 1993; Blank et al. 1998). Die korrelative Unabhängigkeit dieses Faktors ist folgendermaßen zu erklären: Gewalt gegenüber Ausländern wird in Form eines breiten kollektiven Konsenses eher grundsätzlich abgelehnt. Ein spezifischer Zusammenhang zur Identifikation mit der eigenen Nation besteht daher nicht. Der Geschichtsfaktor kann sich aber auch in umgekehrter Weise auswirken: Rechtsorientierte verbinden aufgrund des verlorenen Krieges mit der eigenen Gruppe ein unangenehmes Gefühl. Ihre positive Bewertung kann dagegen durch das Ausblenden der Niederlage motiviert sein (vgl. Mitscherlich/Mitscherlich 1977). Beide Motive können auch Grundlage einer unterschiedlichen Bewertung der deutschen Geschichte insgesamt sein. Das empirische Ergebnis der Nullkorrelation ist deshalb plausibel.

Als Fazit der Faktorenanalyse konnten zwei Dimensionen des Peinlichkeitsgefühls in bezug auf nationale Identität getrennt werden. Zwei weitere Dimensionen zeichnen sich aufgrund der Ergebnisse ab, sind aber meßtheoretisch noch nicht befriedigend operationalisiert.

5. Zum Zusammenhang der Peinlichkeitsbewertungen mit Stolz auf kollektive Aspekte, Fremdgruppenabwertung, nationale Identität und Antisemitismus

Zur Prüfung der externen inhaltlichen Gültigkeit der vier latenten Faktoren verwenden wir folgende Indikatoren und Skalen aus der Münsteraner Stichprobe (MS-II): die Nationalstolzskala (vgl. Blank/Schmidt 1993), die allgemeine Fremdgruppenabwertungsskala (vgl. Blank/Schwarzer 1994; Blank/Wasmer 1996), die Patriotismus-Skala von Kosterman/Feshbach (1989) zur Messung der Bedeutsamkeit der Gruppenzugehörigkeit (Blank/Schmidt 1997; Blank/Wittenberg/Schneider 1998) und Antisemitismus-Indikatoren von Bergmann/Erb (1991). Für die Konzepte Fremdgruppenabwertung, nationale Identität und Antisemitismus sind die verwendeten Indikatoren in Tabelle 4 dargestellt. Jede Skala ist für sich genommen faktorenanalytisch eindimensional.

Für den Stolz auf kollektive Aspekte haben wir hier auf die Darstellung der Itemformulierungen und Häufigkeitsverteilungen verzichtet. Die Itemformulierungen dieser elf Indikatoren sind bei Blank/Schmidt (1993) dokumentiert. Bivariate Korrelationsanalysen zwischen diesen Stolzindikatoren und den Peinlichkeitsitems zeigen, daß von 132 möglichen Korrelationen lediglich 55, das sind 41,7 Prozent, überhaupt signifikant waren (bei $N=308$). 45 dieser signifikanten Korrelationen sind kleiner als .30, was für unsere Annahme einer relativen Unabhängigkeit zwischen positiven und negativen Emotionen spricht. Die numerisch größte Korrelation (.44) besteht zwischen dem Stolz auf die deutsche Geschichte und der Peinlichkeit der Niederlage 1945: Mit zunehmendem Geschichtsstolz nimmt auch die Peinlichkeit über die Niederlage zu. Aufgrund dieser schwachen korrelativen Beziehungen zwischen Stolz und Peinlichkeit verzichten wir im folgenden auf die weitere systematische Berücksichtigung der Stolzemotionen.

Das oben dargestellte vierfaktorielle Meßmodell der Peinlichkeit wurde deshalb um die drei Faktoren Fremdgruppenabwertung, nationale Identität und Antisemitismus erweitert. Unsere Validierungsannahmen lauten:

- der Peinlichkeitsfaktor P-1 ist inhaltlich negativ mit Fremdgruppenabwertung,
- der Peinlichkeitsfaktor P-2 ist inhaltlich negativ mit nationaler Identität,
- der Peinlichkeitsfaktor P-3 ist inhaltlich negativ und der Peinlichkeitsfaktor P-4 inhaltlich positiv mit Antisemitismus korreliert.

Die empirische Überprüfung ergab das in Tabelle 5 dargestellte Meßmodell.¹¹⁾ Auch hier werden die standardisierten Faktorenladungen als formale Validitätskoeffizienten angesehen (vgl. Bollen 1989).

Tabelle 4: Itemformulierungen, Mittelwerte und Standardabweichungen der externen Kriterien (MS-II, N = 308)

FREMDGRUPPENABWERTUNG			
Skalierung: 1 = „stimme überhaupt nicht zu“ bis 7 = „stimme voll und ganz zu“		x	s
Ausl_Anpassg	Die in der Bundesrepublik lebenden Ausländer sollten ihren Lebensstil ein bißchen besser an den der Deutschen anpassen.	3.20	2.06
Ausl_Politik	Man sollte den in der Bundesrepublik lebenden Ausländern jede politische Betätigung in Deutschland untersagen.	2.12	1.82
Ausl_Arbeit	Wenn Arbeitsplätze knapp werden, sollte man die in der Bundesrepublik lebenden Ausländer wieder in ihre Heimat zurückschicken.	1.74	1.53
Ausl_Ehe	Die in der Bundesrepublik lebenden Ausländer sollten ihre Ehepartner unter ihren eigenen Landsleuten auswählen.	1.47	1.34
BEDEUTSAMKEIT DER GRUPPENZUGEHÖRIGKEIT			
		x	s
Bürger	Wie wichtig ist für Sie die Tatsache, daß Sie Bundesbürger/in sind? 1 „überhaupt nicht wichtig“ - 7 „sehr wichtig“.	4.20	2.18
Bindung_Dtld	Wieviel bedeutet es Ihnen, eine innere Bindung zu Deutschland zu haben? 1 „überhaupt nichts“ - 7 „sehr viel“.	4.27	2.00
Vaterland	Wie sehr lieben Sie Ihr Vaterland? 1 „überhaupt nicht“ - 7 „sehr“.	4.20	2.07
ANTISEMITISMUS			
Skalierung: 1 = „stimme überhaupt nicht zu“ bis 7 = „stimme voll und ganz zu“		x	s
Unheil	Juden haben in der Geschichte viel Unheil gestiftet.	1.71	1.45
Jud_Schuld	Durch ihr Verhalten sind die Juden an ihren Verfolgungen nicht ganz unschuldig.	1.92	1.63
Jud_Einfluß	Juden haben auf der Welt zuviel Einfluß.	1.92	1.61

Tabelle 5 zeigt, daß gegenüber der internen Validierung (vgl. Tabelle 2) die Zuordnung der Peinlichkeitsindikatoren zu den Faktoren P-1 und P-2 stabil geblieben sind. Auch die Rangfolge der Faktorenladungen ist in beiden Modellen annähernd gleich. Bei den Faktoren P-3 und P-4 wurde durch die Einführung der externen Kriterien eine höhere

Trennschärfe im Hinblick auf die Indikatoren erreicht. So besteht P-3 nunmehr allein aus dem Indikator dt.Geschichte, während *Niederlg.1945* jetzt eindeutig dem Faktor P-4 zugeordnet ist. Zumindest bei den beiden Faktoren P-1 und P-2 ist deshalb von einer hohen internen Gültigkeit auszugehen.

Tabelle 5: Vollstandardisierte Faktorenladungen der Peinlichkeitsindikatoren und der externen Kriterien

	Antisemitismus	Fremdgruppenabwertung	Bedeutsamkeit der Gruppenzugehörigkeit	Peinlichkeits-Faktoren				Meßfehler
				P-1	P-2	P-3	P-4	
Unheil	.69							.53
Jud_Schuld	.84							.30
Jud_Einfluß	.63							.60
Ausl_Anpassg		.77						.41
Ausl_Politik		.61		-.30				.62
Ausl_Arbeit		.61						.53
Ausl_Ehe		.57						.68
Bindung_Dtld			.88					.22
Bürger			.79				.22	.38
Vaterland			.68					.35
dt.Tourist				.43				.82
Jud_Verfolg				.73	.17			.47
Gewalt_gg.Ausl				.68				.54
Rechtsradikal				.67				.55
Ausl_Feindschft				.59				.63
Ordnngssinn			-.20	.25	.20			.84
dt.angesehen					.78			.39
dt.Herkunft					.95			.09
Dt.-Sein					.87			.25
dt.Geschichte						1.00		.00
Niederlg.1945							.62	.62
Umgang_NS							.59	.66

Chi-Quadrat = 207,5; df = 183, p = .10, GFI = .94.

Die drei externen Kriterien sind aus meßtheoretischer Sicht befriedigend als eigenständige Faktoren abgebildet. Die durchaus plausiblen Fremdladungen einzelner Indikatoren dürften darauf zurückzuführen sein, daß bei den externen Konstrukten Kurzskalen verwendet wurden. Es liegen somit sieben inhaltlich bestimmbare Faktoren vor, deren Korrelationen (vgl. Tabelle 6) zur Beurteilung der externen Gültigkeit der Peinlichkeitsfaktoren herangezogen werden.

Tabelle 6: Vollstandardisierte Faktorenkorrelationen der Peinlichkeitsdimensionen mit den externen Kriterien

	Antisemitismus	Fremdgruppenabwertung	Bedeutsamkeit der Gruppenzugehörigkeit	P-1	P-2	P-3	P-4
Antisemitismus	1.00						
Fremdgruppenabwertung	.50	1.00					
Bedeutsamkeit der Gruppenzugehörigkeit	.24	.58	1.00				
P-1	-.25	--	--	1.00			
P-2	--	-.13	-.34	--	1.00		
P-3	-.17	-.14	-.23	.45	.32	1.00	
P-4	.56	.76	.43	--	--	--	1.00

Die drei externen Kriterien Antisemitismus, Fremdgruppenabwertung und Bedeutsamkeit der Gruppenzugehörigkeit korrelieren positiv miteinander. Wer antisemitisch eingestellt ist, äußert eher fremdenfeindliche Einstellungen und bewertet seine nationale Zugehörigkeit eher als wichtig. Dieser Zusammenhang entspricht sowohl dem Ethnozentrismuskonzept von Levinson (1950) wie auch der Theorie sozialer Identität von Tajfel (1982, vgl. Tajfel/Turner 1986) und ist insofern plausibel.

Die Korrelationsstruktur der vier Peinlichkeitsfaktoren untereinander entspricht den oben dargestellten Ergebnissen der internen Validierung. Wir sehen darin ein weiteres Indiz für die interne Gültigkeit unserer Operationalisierung.

Aus der Betrachtung der Korrelationsstruktur zwischen Peinlichkeitsfaktoren und externen Einstellungskonstrukten folgt, daß das Peinlichkeitsgefühl bzgl. der Gewalt gegen Ausländer (P-1) nur mit Antisemitismus korreliert. Je stärker gewalttätige

Verhaltensweisen negativ bewertet werden, desto weniger ist man antisemitisch eingestellt. Bedeutsam ist in diesem Zusammenhang, daß P-1 weder mit den anderen Faktoren noch mit Alter, Bildung und politischer Orientierung korreliert ist. Dies verweist auf einen breiten Peinlichkeitskonsens über verschiedene Bevölkerungsschichten hinweg.

Auf den ersten Blick scheint ein Widerspruch darin zu bestehen, daß die Peinlichkeit derartiger Gewalttätigkeiten unabhängig von der Fremdgruppenabwertung ist. Hier muß jedoch genau differenziert werden. Es wird nur Gewalt gegenüber Ausländern als unangenehm empfunden. Andere Einstellungen gegenüber Fremdgruppen bleiben davon unberührt. Zwei Interpretationsmöglichkeiten bieten sich an:

a) Ausländer werden als eine bedeutsame Fremdgruppe wahrgenommen. Daraus resultieren ablehnende Einstellungen ihnen gegenüber (vgl. Levinson 1950, Tajfel 1982, Tajfel/Turner 1986). Juden werden im Gegensatz dazu nicht als bedeutsame Fremd- bzw. Minderheitengruppe begriffen. Statt dessen erinnern die auf sie bezogenen Items im Sinne des „sekundären Antisemitismus“ (Freyhold 1971) an die historische Schuld der Deutschen, eine Minderheit physisch nahezu ausgelöscht zu haben. Dieses Gefühl wird nicht durch eine aktuelle Bedeutsamkeit dieser Gruppe überlagert. Somit kann aus den Faktorkorrelationen gefolgert werden, daß der Faktor P-1 die Zurückweisung von Gewalttätigkeit in einem generellen Sinne abbildet. Gewalt wird im Rahmen eines gesellschaftlichen Konsenses zurückgewiesen, unabhängig davon, gegen wen sie sich richtet. Dafür sprechen auch die Nullkorrelationen der P-1-Indikatoren mit Alter, Bildung und politischer Orientierung.

b) Der Völkermord an den Juden richtete sich gegen alle Mitglieder dieser Gruppe. Dagegen betrifft die aktuell ausgeübte physische Gewalt gegen Ausländer nur einzelne Mitglieder von diesen. Die Abwertung von Fremdgruppen erfolgt heute überwiegend durch gewaltfreiere Methoden der sozialen Isolierung und Desintegration und auf der Einstellungsebene (Bergmann/Erb 1991). Insofern geht die Ablehnung von Gewalt zwar mit der Ablehnung antisemitischer Einstellungen einher, doch werden Ausländer damit noch lange nicht akzeptiert. Schließlich zeigen die vorliegenden Daten ähnlich wie andere Studien (vgl. Blank/Schmidt 1993; SINUS 1980), daß Antisemitismus wie auch die Abwertung von Fremdgruppen besonders bei politisch eher Rechtsorientierten, bei Älteren und bei Personen mit niedrigem Bildungsniveau auftritt.

Der Faktor P-2 „Peinlichkeit des Deutschseins“ korreliert am stärksten mit der Bedeutsamkeit der Gruppenzugehörigkeit. Denjenigen, denen die nationale Zugehörigkeit wichtig ist, ist das Deutschsein nicht peinlich. Zwei Erklärungsmöglichkeiten bieten sich an:

a) Ein auf die Eigengruppe bezogenes generalisiertes Peinlichkeitsgefühl hängt eng mit einer abnehmenden Identifikation mit ihr zusammen. Die Eigengruppe verliert an Bedeutung für das individuelle Selbstkonzept (vgl. Tajfel 1982).

b) Das allgemeine Peinlichkeitsgefühl ist Ausdruck einer verhinderten vorbehaltlosen Identifikation mit der eigenen Nation. Damit läßt es sich im Sinne eines eher patriotischen nationalen Selbstverständnisses interpretieren (Staub 1997). Ein nationalistisches Selbstverständnis zeichnet sich im Gegensatz dazu durch eine vorbehaltlose positive Bewertung der Eigengruppe aus (vgl. Levinson 1950; Blank/Schmidt 1997).

Diejenigen, denen das Deutschsein peinlich ist, empfinden auch die deutsche Geschichte (P-3) als eher negativ. Daß es hierbei stärker um das interne Gruppenselbstverständnis der Deutschen geht und weniger um Aspekte des Umgangs mit Fremdgruppen, zeigt sich an der deutlich schwächeren Korrelation dieses Faktors P-2 mit der Fremdgruppenabwertung und einer Nullkorrelation mit dem Antisemitismus.

Der Faktor P-3 (Peinlichkeit der deutschen Geschichte) enthält zwei Aspekte. Zum einen kann die Geschichte wegen der Shoah als unangenehm im Sinne einer Vergangenheitsbewältigung empfunden werden, zum anderen aber eben auch wegen der Niederlage 1945 im Sinne einer Vergangenheitsverdrängung. Die Korrelationen des Faktors P-3 mit Faktor P-4 und auch der politischen Orientierung sind deshalb nicht signifikant. Sowohl die politische Linke wie auch die politische Rechte haben hier ein ähnliches Empfinden, aber aus gegensätzlichen Gründen. Die externen Korrelationen dieses Faktors sind dagegen durchgängig negativ. Wem die deutsche Geschichte unangenehm ist, der ist eher nicht antisemitisch und fremdenfeindlich eingestellt. Zudem ist für ihn die Gruppenzugehörigkeit nicht so wichtig wie für diejenigen, die keine Peinlichkeit im Bezug auf die deutsche Geschichte empfinden. Die niedrigen externen Korrelationen sprechen dafür, daß bei dieser Peinlichkeitsbewertung die beiden oben genannten Bewertungsaspekte eine Rolle spielen. Hier ist über eine Verbesserung der Operationalisierung und über Subgruppenanalysen nachzudenken.

Der Faktor P-4 (Peinlichkeitsempfinden politisch Rechtsorientierter) korreliert positiv mit Antisemitismus, der Fremdgruppenabwertung und der Bedeutsamkeit der Gruppenzugehörigkeit. Wer die Niederlage sowie die Tatsache, daß immer noch viel über den Nationalsozialismus geredet wird, als negativ empfindet, äußert auch eher antisemitische und fremdenfeindliche Einstellungen. Für ihn ist weiterhin die nationale Zugehörigkeit eher wichtig.

6. Zusammenfassung und Interpretation

Bei den bislang vorliegenden empirischen Studien zur nationalen Identität wurde den der Nation gegenüber bestehenden Emotionen lediglich in einer einzigen Dimension Rechnung getragen. Die bisherige Beschränkung auf den Nationalstolz läßt den Verdacht entstehen, bei der Nation handele es sich um eine sakrosankte Institution (vgl. Theweleit 1978). Demgegenüber liegt es bei einer eingehenden Betrachtung auf der Hand, daß die Zugehörigkeit zu einer Nation, wie bei anderen Gruppenmitgliedschaften auch, mit ambivalenten Gefühlen einhergehen dürfte. Wir haben deshalb ein erstes Meßinstrument zur Erfassung des auf die Nation bezogenen Peinlichkeitsgefühls entwickelt. Die empirische Überprüfung dieses Instruments in einer regionalen Stichprobe kommt zu zwei zentralen Ergebnissen: Das Peinlichkeitsgefühl gegenüber der eigenen Nation und ihren Teilaspekten ist ein eigenständiger Faktor nationaler Identität und es ist multidimensional.

Trotz meßtheoretischer Probleme lassen sich aufgrund der stabilen Befragungsergebnisse klare inhaltliche Rückschlüsse ziehen. Das von uns untersuchte Gefühl der Peinlichkeit läßt sich in folgende vier Faktoren untergliedern. Als unangenehm bewerten die Befragten die Gewalttätigkeit gegenüber Fremdgruppen, die eigene Gruppenzugehörigkeit, die deutsche Geschichte allgemein und die Art des Umgangs mit dem Nationalsozialismus. Dieses Ergebnis verdeutlicht einerseits das Vorhandensein unterschiedlicher inhaltlicher Ebenen, auf die sich das Peinlichkeitsgefühl beziehen kann. Andererseits zeigt sich, daß spezifische Subgruppen auch ein spezifisches Peinlichkeitsempfinden besitzen können.

Ein weiteres Ergebnis betrifft die Beziehung zwischen positiven (Stolz auf kollektive Güter) und negativen Emotionen (Peinlichkeit). Bei dem zweitgenannten Gefühl handelt es sich nicht um das invertierte Äquivalent des Nationalstolzes. Beide Emotionen treten eher unabhängig voneinander auf.

Als Fazit für die Diskussion über nationale Identität ist festzuhalten, daß negative Emotionen als ein relevanter Faktor in die Debatte miteinzubeziehen sind. Nur auf diesem Wege wird eine Annäherung an die Ambivalenzen individueller Gruppenbindung möglich. So zeichnen sich Nationalisten durch die Idealisierung der Eigengruppe aus, was Ablehnung von Ambiguitäten impliziert. Demgegenüber können für Patrioten negative Emotionen Ausdruck eines kritischen Verhältnisses zur Nation im Sinne eines konstruktiven Patriotismus sein (vgl. Staub 1997; Blank/Schmidt 1997).

In weiteren Forschungsschritten wird es zukünftig darauf ankommen zu klären, welche weiteren negativen Emotionen zu konzeptualisieren sind, auf welche Weise positive und

negative Gefühle auf der individuellen Ebene integriert werden und welche individuellen und sozialen Rahmenbedingungen zur Wahrnehmung spezifischer Emotionen gegeben sein müssen. Im Hinblick auf die Beschreibung und Erklärung unterschiedlicher Formen der Systemunterstützung bedarf es einer systematischen theoretischen Erweiterung bestehender Ansätze (z.B. Easton 1975; Tajfel/Turner 1986) und deren empirischer Prüfung.

Korrespondenzadressen

Thomas Blank, M.A.
Westfälische Wilhelms-Universität
Institut für Soziologie, Abt. II
Scharnhorststr. 121
48151 Münster
Tel.: 0251-8329338 oder 02545-1577
email: blankt@uni-muenster.de

Dr. Horst-Alfred Heinrich,
Justus-Liebig-Universität Gießen
Institut für Politikwissenschaften
Karl-Glöckner-Str. 21E
35394 Gießen
Tel.: 0641-67883
email: horst-alfred.heinrich@sowi.uni-giessen.de

Anmerkungen

- 1) Die vorliegenden Ergebnisse entstammen dem Projekt „Nationale Identität der Deutschen“, welches im Rahmen des Schwerpunktprogramms „Sozialer und politischer Wandel im Zuge der Integration der DDR-Gesellschaft“ seitens der DFG von 1992 bis 1998 gefördert wurde (Projekt-Nr.: Schm 658/4-3). Wir danken den Mitgliedern der Projektgruppe Peter Schmidt, Marc Hübner und Stefan Schwarzer für die kritische Durchsicht des Manuskripts sowie den anonymen Gutachtern für wertvolle Anregungen.
- 2) Die Begriffe „Kollektiv“ und „Nation“ verwenden wir synonym. Präzise ausgedrückt ist die Nation eine Form von kollektivem Akteur (vgl. Olson 1968; Coleman 1990).
- 3) Laut Skala (1994) gaben 33% der befragten Deutschen an, sie hätten sich persönlich schon einmal vom Auftreten von Landsleuten im Ausland peinlich berührt gefühlt.
- 4) Auf eine dritte Möglichkeit der Umbewertung der sozialen Situation wird hier nicht weiter eingegangen (vgl. Tajfel 1982).
- 5) Der Fragetext lautete: „Was ist Ihnen als Deutsche/r unangenehm oder peinlich? Bitte nennen Sie mir maximal drei Begriffe.“ Die Worte „unangenehm“ und „peinlich“ wurden semantisch als ähnlich betrachtet. Da zudem der Chi²-Test über die einzelnen Kategorien zwischen GI und MS-I, wo nur nach „peinlich“ gefragt wurde, lediglich bei vier von 18 Begriffsklassen zu signifikanten Abweichungen führt, liegt unserer Ansicht nach kein Doppelstimulus vor.
- 6) Mittels Cohens Kappa ließ sich die Reliabilität des Kategorienschemas bestätigen. Die zufallsbereinigte Übereinstimmung lag bei acht *Ratern* zwischen $K = .77$ und $K =$

.87 ($M_e = .82$) (vgl. Asendorpf/Wallbott 1979; Schmidt 1965). Eine Bestätigung der Validität erfolgte mittels t-Tests zwischen den Antworten auf die geschlossenen Items in MS-I mit denen auf die erneut erhobene offene Frage.

7) Eindeutige Zuordnung heißt, daß mindestens sieben Rater in ihrem Zuordnungsurteil bezüglich eines Begriffes übereinstimmen.

8) Polychorische Korrelationen wurden wegen des ordinalen Meßniveaus und den signifikanten Verletzungen der Normalverteilungsannahme in den empirischen Häufigkeitsverteilungen berechnet (vgl. Jöreskog/Sörbom 1988).

9) Als Input verwendeten wir polychorische Korrelationen und die asymptotische Varianz-Kovarianzmatrix. Der WLS-Schätzer liefert in Kombination mit polychorischen Korrelationen die optimalen Schätzwerte (vgl. Jöreskog/Sörbom 1993a). Es sind nur die signifikanten Faktorenladungen dargestellt.

10) Implizit erklärt dieses Resultat auch die Emotionen, mit denen die Debatten über den Nationalsozialismus und seine Bewertung immer wieder ausbrechen.

11) Neben dieser Modellspezifikation mußten acht Meßfehlerkorrelationen zugelassen werden, um eine ausreichende Modellanpassung zu erreichen. Auf eine ausführliche Diskussion der spezifischen Meßqualität unserer Operationalisierung verzichten wir hier, da im Mittelpunkt unserer Untersuchung der Nachweis steht, daß negative Emotionen gegenüber der Eigengruppe ein eigenständiger Aspekt nationaler Identität sind.

Literatur

Adorno, Th. W. et al., 1950: The authoritarian personality. New York: Norton.

Ajzen, I., 1988: Attitudes, personality and behavior. Milton Keynes: Open University Press.

Almond, G. A./Verba, S., 1963: The civic culture. Political attitudes and democracy in five nations. Princeton, NY: Princeton University Press.

Asendorpf, J./Wallbott, H. G., 1979: Maße der Beobachterübereinstimmung: Ein systematischer Vergleich. Zeitschrift für Sozialpsychologie 10: 243-252.

Baring, A., 1992: Deutschland, was nun? Berlin: Siedler.

Bergmann, W./Erb, R., 1991: Antisemitismus in der Bundesrepublik Deutschland. Ergebnisse der empirischen Forschung von 1946-1989. Opladen: Leske & Budrich.

Blank, Th., 1994: Pretest 4: Politbarometer 2/93 Münster. Stichprobenbeschreibung, Fragebogen und Häufigkeitsauszählung. Heft 5 der Schriftenreihe Nationale Identität. Arbeitsberichte aus dem DFG-Projekt „Nationale Identität der Deutschen. Messung und Erklärung der Veränderungsprozesse in Ost und West“. Gießen: ohne Verlag.

- Blank, Th., 1997: Wer sind die Deutschen? Nationalismus, Patriotismus, Identität - Ergebnisse einer empirischen Längsschnittstudie. Aus Politik und Zeitgeschichte B13/97, 21. März 1997: 38-46.
- Blank, Th. et al., 1993: Pretest 1: Politbarometer 1/92 Gießen. Stichprobenbeschreibung, Fragebogen und Häufigkeitsauszählung. Heft 1 der Schriftenreihe Nationale Identität. Arbeitsberichte aus dem DFG-Projekt „Nationale Identität der Deutschen. Messung und Erklärung der Veränderungsprozesse in Ost und West“. Gießen: ohne Verlag.
- Blank, Th. et al., 1994: Pretest 5: Politbarometer 1/94 Münster. Stichprobenbeschreibung, Fragebogen und Häufigkeitsauszählung. Heft 6 der Schriftenreihe Nationale Identität. Arbeitsberichte aus dem DFG-Projekt „Nationale Identität der Deutschen. Messung und Erklärung der Veränderungsprozesse in Ost und West“. Gießen: ohne Verlag.
- Blank, Th./Wittenberg, J./Schneider, H., 1998: Constructive patriotism: The concept and its measurement. *Social Indicators Research* (im Druck).
- Blank, Th./Schmidt, P., 1993: Verletzte oder verletzende Nation? Empirische Befunde zum Stolz auf Deutschland. *Journal für Sozialforschung* 33 (4): 49-73.
- Blank, Th./Schmidt, P., 1997: Konstruktiver Patriotismus im vereinigten Deutschland? Ergebnisse einer repräsentativen Studie. S. 127-148 in: A. Mummendey/B. Simon (Hrsg.), *Identität und Verschiedenheit*, Bern: Huber.
- Blank, Th./Schwarzer, S., 1994: Ist die Gastarbeiterskala noch zeitgemäß? Die Reformulierung einer ALLBUS-Skala. *ZUMA-Nachrichten* 34: 97-115.
- Blank, Th./Wasmer, M., 1996: Gastarbeiter oder Ausländer? Ergebnisse des Splits mit den reformulierten Gastarbeiterfragen im ALLBUS 1994. *ZUMA-Nachrichten* 38: 45-69.
- Bollen, K. A., 1989: *Structural equations with latent variables*. New York: Wiley.
- Bubis, I./Schäuble, W., 1996: *Deutschland wohin?* Hrsg. von Frank Schirrmacher. Herder.
- Coleman, J. S., 1990: *Foundations of social theory*. Cambridge/Mass.: Harvard University Press.
- Eagly, A. H./Chaiken, S., 1993: *The Psychology of attitudes*. San Diego: Harcourt Brace Janovich.
- Easton, D., 1975: A Re-Assessment of the Concepts of Political Support. *British Journal of Political Science* 5: 435-457
- Fishbein, M./Ajzen, I., 1975: *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading/Mass.: Addison-Wesley.
- Fiske, S. T./Pavelchak, M. A., 1986: Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. S. 167-203 in R. M. Soerentino/E.

- T. Higgins (Hrsg.), *Handbook of motivation and cognition: Foundations of social behavior*. New York: Guilford Press.
- Freyhold, M. v., 1971: *Autoritarismus und politische Apathie. Analyse einer Skala zur Ermittlung autoritätsgebundener Verhaltensweisen*. Frankfurt.
- Giesen, B./Leggewie, C. (Hrsg.), 1991: *Experiment Vereinigung. Ein sozialer Großversuch*. Berlin: Rotbuch.
- Giordano, R., 1987: *Die zweite Schuld oder Von der Last Deutscher zu sein*. Hamburg: Rasch & Röhrig.
- Herdegen, G., 1987: *Einstellungen der Deutschen (West) zur nationalen Identität. Politische Vierteljahresschrift, Sonderband 18: 205-221*.
- Hirschman, A. O., 1970: *Exit, voice, and loyalty: responses to decline in firms, organizations, and states*. Cambridge: Harvard University Press.
- Jöreskog, K. G./Sörbom, D., 1988: *PRELIS. A Program for Multivariate Data Screening and Data Summarization. A Preprocessor for LISREL. 2nd Edition*. Mooresville: Scientific Software.
- Jöreskog, K. G./Sörbom, D., 1993a: *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Hillsdale, Hove, London: Erlbaum.
- Jöreskog, K. G./Sörbom, D., 1993b: *PRELIS 2. User's reference guide*. Chicago: SPSS Inc.
- Kosterman, R./Feshbach, S., 1989: *Toward a measure of patriotic and nationalistic attitudes*. *Political Psychology* 10: 257-274.
- Kruse, O., 1985: *Emotionsdynamik und Psychotherapie*. Weinheim: Beltz.
- Levinson, D., 1950: *The study of ethnocentric ideology*. S. 102-150 in: Th. W. Adorno et al., *The authoritarian personality*. New York: Norton.
- Lüdemann, C., 1995: *Fremdenfeindliche Gewalt und Lichterketten. Kollektives Handeln als „Rational Choice“*. S. 355-381 in: Lederer, G./Schmidt, P. (Hrsg.), *Autoritarismus und Gesellschaft. Trendanalysen und vergleichende Jugenduntersuchungen 1945-1993*. Opladen: Leske+Budrich
- Merten, K., 1983: *Inhaltsanalyse. Einführung in Theorie, Methode und Praxis*. Opladen: WV.
- Merton, R. K., 1949: *Social structure and anomie*. S. 185-214 in: R. K. Merton (Hrsg.), *Social theory and social structure. (enlarged edition)*, New York: Free Press.
- Mitscherlich, A./Mitscherlich, M., 1967: *Die Unfähigkeit zu trauern*. München: Piper.
- Mohler, P. Ph./Götze, H., 1992: *Worauf sind die Deutschen stolz? Eine vergleichende Analyse zur gefühlsmäßigen Bindung an das politische System der Bundesrepublik*. S. 45-63 in: Mohler, P. Ph./Bandilla, W. (Hrsg.): *Blickpunkt Gesellschaft 2: Einstellungen und Verhalten der Bundesbürger in Ost und West*. Opladen: Westdeutscher Verlag.

- Mummendey, H. D., 1992: Eine Skala zum deutschen Nationalstolz. Bielefelder Arbeiten zur Sozialpsychologie, Nr. 163, Bielefeld: Universität Bielefeld.
- Noelle-Neumann, E./Köcher, R., 1987: Die verletzte Nation. Über den Versuch der Deutschen, ihren Charakter zu ändern. Stuttgart: Deutsche Verlagsanstalt.
- Olson, M. E., 1968: Die Logik des kollektiven Handelns und die Theorie der Gruppen. Tübingen: Mohr.
- Ottati, V. C./Wyer, R. S. Jr., 1993: Affect and political judgment. S. 296-315 in: S. Iyengar/W. J. McGuire (Hrsg.), Explorations in political psychology. Durham, London: Duke University Press.
- Rahn, W. M. et al., 1994: A framework for the study of public mood. Unpublished paper.
- Reinecke, J., 1991: Interviewer und Befragtenverhalten: Theoretische Ansätze und methodische Konzepte. Opladen: Westdeutscher Verlag.
- Rosenberg, M. J./Hovland, C. I., 1960: Cognitive, affective and behavioral components of attitude. S. 1-14 in: M. J. Rosenberg et al. (Hrsg.), Attitude organization and change: An analysis of consistency among attitude components. New Haven.
- Schmidt, H. D., 1965: Die Beurteilung des menschlichen Verhaltens durch Ratingskalen. Diss., Bonn.
- Schweigler, G., 1973: Nationalbewußtsein in der BRD und der DDR. Düsseldorf: Universitätsverlag.
- SINUS, 1980: Rechtsextreme politische Einstellungen in der Bundesrepublik Deutschland. Abschlußbericht. Heidelberg/München.
- Skala, 1994: Dokumentation einer Infas-Umfrage für das ARD-Fernsehen. Frankfurter Rundschau 50, Nr. 173: 18.
- Staub, E., 1997: Blind versus constructive patriotism: Moving from embeddedness in the group to critical loyalty and action. S. 213-228 in: D. Bar-Tal/E. Staub (Hrsg.), Patriotism in the lives of individuals and nations, Chicago: Nelson-Hall.
- Tajfel, H., 1982: Gruppenkonflikt und Vorurteil. Entstehung und Funktion sozialer Vorurteile. Bern, Stuttgart, Wien: Huber.
- Tajfel, H./Turner, J. C., 1986: The social identity theory of intergroup behavior. S. 7-24 in: S. Worchel/W. G. Austin (Hrsg.), Psychology of intergroup relations. Chicago: Nelson-Hall.
- Theweleit, K., 1978: Männerphantasien. Bd. 2. Frankfurt: Roter Stern.
- Topf, R. et al., 1990: Nationalstolz in Großbritannien und der Bundesrepublik Deutschland. S. 172-190 in: W. Müller et al. (Hrsg.), Blickpunkt Gesellschaft. Einstellungen und Verhalten der Bundesbürger. Opladen: Westdeutscher Verlag.

Vester, H.-G., 1991: Emotion, Gesellschaft und Kultur. Grundzüge einer soziologischen Theorie der Emotionen. Opladen: Westdeutscher Verlag.

Wakenhut, R., 1994: Die deutsche Nation im Bewußtsein junger Bundesbürger. Begriffliche Klärungen und empirische Befunde zum Bewußtsein nationaler Zugehörigkeit in Deutschland. Eichstätt: unveröff. Manuskript.

Westle, B., 1994: Traditionalismus, Verfassungspatriotismus und Postnationalismus im vereinigten Deutschland. S. 43-76 in Niedermayer, O./Beyme, K. von (Hrsg.), Politische Kultur in Ost- und Westdeutschland. Berlin: Akademie Verlag.

Zanna, M. P./Rempel, J. K., 1988: Attitudes: a new look at an old concept. S. 315-334 in: D. Bar-Tal/A. W. Kruglanski (Hrsg.), The social psychology of knowledge. New York.

Zitelmann, R., 1990: Die Deutsche Frage. Analysen und Standpunkte. Zeitschrift für Politik N.F. 37: 322-349.

ZEITBEWUßTSEIN, HANDLUNGSINTENTIONEN UND EIGENVERANTWORTUNG: EINE ANALYSE DER ZUSAMMENHÄNGE ZWISCHEN HANDLUNGSBESTIMMENDEN ORIENTIERUNGEN*)

HELGA SIEVERS

Zeit wird von den Menschen auf unterschiedliche Art wahrgenommen und interpretiert. Je nachdem, wie sie die Zusammenhänge zwischen Vergangenheit, Gegenwart und Zukunft sowie die Bedeutung ihres eigenen Handelns im Zeitablauf interpretieren, ergeben sich unterschiedliche Implikationen hinsichtlich ihrer Handlungsintentionen. Menschen können ihr Leben aktiv gestalten oder passiv abwarten, was passieren wird. Sie können danach streben, ihren eigenen Nutzen zu maximieren oder sich für das Wohlergehen anderer Menschen einsetzen. Die Ausgestaltung der Handlungsintentionen wird dabei in entscheidendem Umfang von ihrem Zeitverständnis beeinflusst. Eine weitere Determinante des Verhaltens ist die Wahrnehmung von Eigenverantwortung. Je größer das Ausmaß an wahrgenommener Eigenverantwortung ist, desto aktiver wird das eigene Leben gestaltet. Dabei wird die Wahrnehmung von Eigenverantwortung ebenfalls von der Form des Zeitbewußtseins beeinflusst. In diesem Aufsatz soll gezeigt werden, wie sich die Zusammenhänge zwischen dem Zeitverständnis, aktiven vs. passiven sowie individualistischen vs. kollektiven Handlungsorientierungen sowie der Perzeption von Eigenverantwortung gestalten. Mit Hilfe von verschiedenen Indikatoren soll hier eine empirische Überprüfung der vorgestellten Hypothesen geleistet werden. Abschließend soll mittels einer multivariaten Analyse ein komplexes Erklärungsmodell der dargelegten Zusammenhänge erstellt werden.

Time is perceived and interpreted by people in different ways. The way they interpret the connections between the past, the present and the future and the meaning of their own behaviour in time has different implications for their intended behaviour. People may organize their life in an active or passive way. They may strive to maximize their own interests or, alternatively, those of other people. The organisation of intentions is strongly influenced by time perspectives. Another determinant of behaviour is the

perception of self-responsibility. The greater the degree of perception of self-responsibility, the more actively people structure their own lives. Thus perceptions of self-responsibility are also influenced by time perspectives. The article examines the character of the connections between time perspective, active vs. passive as well as individualistic vs. collective orientations of behaviour and the perception of self-responsibility. Several indicators are used to test the hypotheses presented empirically. Finally a complex model is constructed to explain the demonstrated relationships.

1. Zeitbewußtsein als Erklärungsvariable für das Ausmaß an Handlungsbereitschaft

Die Zeitsoziologie geht von der Annahme aus, daß Menschen Zeit unterschiedlich wahrnehmen und unterschiedlich mit ihr umgehen, sie also verschiedenen Formen des „Zeitbewußtseins“ anhängen. Im Mittelpunkt dieser zeittheoretischen Konzeption steht die Frage, wie der Zusammenhang zwischen Vergangenheit, Gegenwart und Zukunft individuell wahrgenommen und interpretiert wird. Von Bedeutung ist dabei, inwieweit die Gegenwart von der Vergangenheit beeinflusst ist, und wie die Zukunft durch das gegenwärtige Handeln gesteuert werden kann. Es geht hier somit um die subjektive Interpretation von Veränderungen und Ereignissen in der Zeit sowie um die Bedeutung des eigenen Handelns (Häder/Mohler 1995: 20; Häder 1996: 18). Das Zeitbewußtsein dient der Ordnung der wahrgenommenen Veränderungen. Die Perzeption des Zusammenhangs von Vergangenheit, Gegenwart und Zukunft und die Wahrnehmung der eigenen Position im Zeitablauf hat einen entscheidenden Einfluß auf die Entwicklung von Handlungsintentionen und damit auf das Verhalten. So ist es für Luhmann unbestritten, daß „... der Handelnde nach seinen Intentionen handelt und daß seine Zeitvorstellung die Struktur seiner Intentionen bestimmt, etwa die Weite seines Zeithorizontes, seine Risikobereitschaft, seine Bereitschaft zur Vertagung von Befriedigungen“ (Luhmann 1979: 63). Entsprechend findet sich eine Vielzahl an empirischen Arbeiten, die einen Zusammenhang zwischen der jeweiligen Zeitperspektive und bestimmten Verhaltensmustern herstellen (vgl. auch Häder 1996: 40).

Im folgenden soll ein Indikator zum Zeitbewußtsein verwendet werden, der von Häder und Mohler (1995) entwickelt wurde. Diese Indikator wurde in Anlehnung an die Arbeiten von Fischer (1981), Wendorff (1988) und insbesondere Rammstedt (1975) entwickelt, die vier verschiedene Formen des Zeitbewußtseins unterscheiden:

1. Okkasionales Zeitbewußtsein

Zeit wird nicht als kontinuierlicher Fluß von Vergangenheit - Gegenwart - Zukunft verstanden, sondern als erlebte Folge nichtkontinuierlicher Ereignisse. Damit erfolgen Veränderungen nicht gesetzmäßig und sind somit auch nicht steuerbar. Im Mittelpunkt dieses Zeitverständnisses steht die Gegenwart, das „Hier und Jetzt“. Vergangenheit und Zukunft haben keine Bedeutung, das Handeln wird so strukturiert, daß lediglich dessen aktuelles Moment berücksichtigt wird.

2. Zyklisches Zeitbewußtsein

Dieses Zeitbewußtsein beruht auf der Vorstellung eines kreisförmigen Verlaufs aller Prozesse. Vergangenheit und Zukunft sind für die Gestaltung der Gegenwart ohne Bedeutung, da alles, was passiert, nur eine Wiederholung bereits erlebter Vorgänge darstellt. Die Zeiterfahrung unterscheidet zwar ein Vorher und Nachher, diese unterliegen aber keinem linearen Ablauf und kausalen Gesetzmäßigkeiten. Da die Zukunft prädestiniert ist, kann sie nicht durch das gegenwärtige Handeln beeinflusst werden, eine bewußte Verhaltensplanung wird damit überflüssig.

3. Linear geschlossenes Zeitbewußtsein

Prozesse im Zeitablauf werden erfahren als ein irreversibler, progressiver Ablauf auf ein letztes Ziel hin (z.B. eine religiöse Orientierung oder ein Karriereziel). Vergangenheit, Gegenwart und Zukunft werden streng unterschieden, um diesen Prozeß erfahrbar zu machen, d.h. die Vergangenheit muß als Vorher bewahrt werden, um ihre Bedeutung für die Gegenwart ermessen zu können. Die Zukunft ist planbar, sie kann durch das gegenwärtige Handeln gestaltet werden. Der Schwerpunkt dieses Zeitbewußtseins liegt in der Zukunft.

4. Linear offenes Zeitbewußtsein

Vergangenheit, Gegenwart und Zukunft werden wie beim linear geschlossenen Zeitbewußtsein als lineares Kontinuum verstanden. Die Gegenwart ist das Ergebnis der Vergangenheit, und Zukunft kann durch das gegenwärtige Handeln beeinflusst werden. In dieser Form des Zeitbewußtseins ist aber die Gegenwart zentral; aus der Gegenwart ergeben sich für die Zukunft mehr Möglichkeiten, als verwirklicht werden können. Dieses Potential an Möglichkeiten ist abhängig von der gegenwärtigen Gesellschaftsstruktur, so daß die Zukunft in der Gegenwart durch Selektion beeinflusst wird.¹⁾

Das Zeitverständnis eines Menschen hat u.a. eine Bedeutung für das Ausmaß an Anstrengungen, die er bei der Gestaltung des eigenen Lebens unternimmt. Die einzelnen

Formen des Zeitbewußtseins sind mit jeweils spezifischen Handlungsimplikationen verbunden. Insbesondere die Zukunftsorientierung gilt als der Aspekt der Zeitperspektive, der eine hohe Bedeutung für soziales Handeln und Entscheiden hat. LeShans (1952) geht davon aus, daß eine positive Beziehung zwischen der Zukunftsperspektive und der Fähigkeit des Befriedigungsaufschubs besteht. Das würde bedeuten, daß Menschen, die bei ihrem gegenwärtigen Handeln auch die Implikationen für die Zukunft berücksichtigen, in der Gegenwart bereit sind, bestimmte Einschränkungen in Kauf zu nehmen. So ist zu vermuten, daß Menschen, deren Orientierung in der Zukunft liegt, in der Gegenwart in ihre Bildung und Qualifizierung investieren, um in der Zukunft sozialen Erfolg zu haben.

Die Verknüpfung von Zeitbewußtsein und der Fähigkeit des Befriedigungsaufschubs, die als Voraussetzung für sozialen Erfolg gilt, führt gleichzeitig zu einer Verbindung des Zeitbewußtseins mit einer Reihe von sozialstrukturellen Variablen, wie z.B. Bildung, Alter, soziale Herkunft, Status etc. Die Arbeit von LeShans konnte entsprechend signifikante Unterschiede hinsichtlich der Weite der Zukunftsperspektive, der Frustrationstoleranz und der Verhaltensausrichtung zwischen Unterschicht- und Mittelschichtkindern feststellen. Auch Huinink und Häder (1998) vermuten einen Zusammenhang zwischen Alter und Zeitverständnis. Mit zunehmenden Alter kommt es zu einem Rückgang der aktiven Lebensgestaltung und damit auch zu einer tendenziellen Abkehr von einem linearen Zeitverständnis. Zieht man außerdem die Überlegung in Betracht, daß mit zunehmenden Alter wichtige Entwicklungsprozesse (z.B. in beruflicher oder familialer Hinsicht) weitgehend abgeschlossen sind, dann kann davon ausgegangen werden, daß insbesondere das zyklische Zeitbewußtsein an Bedeutung gewinnt.

Ein weiterer denkbarer Zusammenhang könnte zwischen Zeitorientierungen und gesellschaftlichen Zielen, wie sie im Inglehart-Index erfaßt werden, bestehen. Der Inglehart-Index ergibt sich aus der Präferenz für „materialistische“ oder „postmaterialistische“ Ziele. Während die materialistische Ausrichtung auf eine Bewahrung des gesellschaftlichen Status quo gerichtet ist („Aufrechterhaltung von Ruhe und Ordnung“, „Kampf gegen steigende Preise“), orientieren sich die postmaterialistischen Ziele mehr an Partizipation und Selbstverwirklichung („Mehr Einfluß der Bürger auf die Entscheidungen der Regierung“, „Schutz des Rechts auf freie Meinungsäußerung“). Im Zusammenhang mit dem hier vorgestellten Zeitverständnis könnte man vermuten, daß die materialistischen Ziele eher von Menschen präferiert werden, die die Entscheidungen für das gesellschaftliche Wohl eher anderen überlassen. Dies könnte damit zusammenhängen, daß sie selbst nur wenig Einfluß auf ihr Leben wahrnehmen und sich eher als von außen determiniert sehen. In diesem Fall ist von einem

Zusammenhang zwischen der Präferenz für materialistische Ziele und dem zyklischen Zeitbewußtsein auszugehen. Postmaterialistische Werte gehen dagegen von Selbstbestimmung und Einflußnahme des Individuums auf den politischen Prozeß aus. Von daher ist hier ein Zusammenhang zwischen Postmaterialismus und den linearen Zeittypen zu erwarten.

Coser und Coser (1963) verbinden die Zeitperspektive mit weiteren Orientierungsarten, nämlich mit kollektiven vs. individualistischen sowie aktiven vs. passiven Orientierungen. Dieser Ansatz läßt weitere Schlußfolgerungen zu. Während der individualistisch-aktive Typ allein am individuellen Erfolg orientiert ist, strebt der kollektiv-aktive Typ nach anderen Zielen; er findet sein Ziel in der uneigennütigen Arbeit in Organisationen (z.B. Gewerkschaften) oder verfolgt eine utopische Zukunft. Der kollektiv-passive Typ setzt nicht auf aktives Handeln, sondern hofft auf externe Umwälzungen. Der individualistisch-passive Typ verfolgt eine hedonistische Zeitorientierung; er lebt im Hier und Jetzt ohne Bindung an soziale Gruppen.

Setzt man diese vier Orientierungsarten in Verbindung mit den oben vorgestellten Zeittypen, dann lassen sich hieraus interessante Schlüsse ziehen. Wie bereits erwähnt, sieht sich der zukunftsorientierte Zeittyp (linear geschlossen, linear offen) in der Lage, durch sein gegenwärtiges Handeln die Zukunft zu beeinflussen. Es wird von ihm ein *aktives* Handeln erwartet. Nach der Typenbildung von Coser und Coser kann dieses Handeln sich nun individuell (auf den eigenen Nutzen) oder aber sozial (auf den Nutzen anderer) ausrichten. Vom zyklischen und okkasionalen Zeittyp ist dagegen ein eher *passives* Verhalten zu erwarten. Der zyklische Zeittyp erwartet keine grundlegenden Veränderungen, so daß ein aktives Handeln für ihn überflüssig ist. Der okkasionale Zeittyp sieht womöglich schon einen Einfluß, dieser ist aber ohne Bedeutung, da er in der Gegenwart ohne Blick auf die Zukunft lebt.

Eine weitere Überlegung, die hier angestellt werden soll, betrifft den Zusammenhang von Zeitbewußtsein und dem Ausmaß an perzipierter *Eigenverantwortung*. Eigenverantwortung wird hier als Persönlichkeitsmerkmal verstanden, d.h. als eine stabile Konfiguration von Zielvorstellungen, Motiven und Plänen, in deren Mittelpunkt die Verantwortung für sich selbst und das eigene Leben als zentrale motivationale Größe steht. Verantwortung in diesem Sinne ist somit immer im Rahmen einer Handlungstheorie zu verordnen (Bierhoff 1994). Sowohl das linear geschlossene als auch das linear offene Zeitverständnis gehen davon aus, daß die Entwicklung der Zukunft durch das eigene Handeln gestaltet werden kann. Der zyklische Zeittyp sieht sich dagegen eher als Opfer von Entwicklungen, auf die er keinen Einfluß nehmen kann. Diese beiden Formen des Zeitverständnis unterscheiden sich somit hinsichtlich des

Ausmaßes an wahrgenommener Entwicklungskontrolle. Hier könnte man nun vermuten, daß diese Unterschiede auch eine Wirkung auf die Bereitschaft haben, für das eigene Leben die Verantwortung zu übernehmen. Zu erwarten ist, daß die linearen Zeittypen, die einen Einfluß auf ihr Leben wahrnehmen, auch ein hohes Maß an Eigenverantwortung aufweisen. Der zyklische Zeittyp hingegen empfindet sich als eher fremdgesteuert und schiebt die Verantwortung für sein Leben auf andere Kräfte. Er wird wahrscheinlich nur ein geringes Maß an Eigenverantwortung empfinden. Für den okkasionalen Zeittypen kann hier wiederum keine Aussage getroffen werden. Dieser Zeittyp ist dadurch geprägt, daß er ohne Blick auf die Zukunft lebt, welchen Einfluß er auf sein Leben wahrnimmt, ist nicht zu sagen.

Faßt man diese Überlegungen zusammen, dann kommt man zu folgenden Hypothesen:

1. Sowohl der linear offene als auch der linear geschlossene Zeittyp orientieren sich an der Zukunft. Sie gestalten ihr Leben aktiv und zeigen entsprechend ein hohes Ausmaß an Handlungsbereitschaft. Diese kann sich sowohl auf den eigenen Nutzen als auch auf den Nutzen anderer richten. Außerdem zeigen diese Zeittypen ein hohes Maß an Eigenverantwortung: sie sehen sich selbst als für ihr eigenes Wohl verantwortlich an.
2. Der zyklische Zeittyp sieht sein Leben als von außen bestimmt an. Er empfindet keinen eigenen Einfluß auf seine Lage und verhält sich daher eher passiv. Es werden keine großen Handlungsanstrengungen von ihm erwartet. Er empfindet sich als fremdgesteuert und wälzt die Verantwortung für sein Leben auf andere ab.
3. Der okkasionale Zeittyp lebt im Hier und Jetzt. Da er nicht an der Zukunft interessiert ist, sind entsprechende Bemühungen, diese zu gestalten, überflüssig. Von ihm sind ebenfalls keine Handlungsanstrengungen zu erwarten. Hinsichtlich des Ausmaßes an Eigenverantwortung kann hier keine Aussage getroffen werden.

2. Welche Faktoren beeinflussen das Zeitbewußtsein?

Im folgenden sollen die oben angestellten Überlegungen empirisch überprüft werden. Es werden die Daten der Erhebung „Leben Ostdeutschland 1993“ verwendet, wo der hier vorgestellte Zeitindikator erstmals verwendet wurde.²⁾

Wie oben bereits dargestellt, ist zu vermuten, daß das Zeitbewußtsein von verschiedenen soziostrukturellen Merkmalen wie Alter und Bildung beeinflusst wird. Es wird außerdem vermutet, daß ein Zusammenhang mit einer bestimmten gesellschaftlichen Werteorientierung vorliegt. Um dies zu überprüfen, wurde jeweils eine Kreuztabellierung mit Chi-Quadrat-Test auf Unabhängigkeit zwischen den Variablen

Zeitbewußtsein mit Alter und Bildung sowie dem Inglehart-Index vorgenommen. Die Zusammenhänge, die hier festgestellt werden konnten, sind sämtlich nach dem Chi-Quadrat-Test auf Unabhängigkeit auf dem 0,001-Niveau signifikant. Es zeigt sich folgendes Bild:

Abbildung 1: Zeitbewußtsein und Bildung

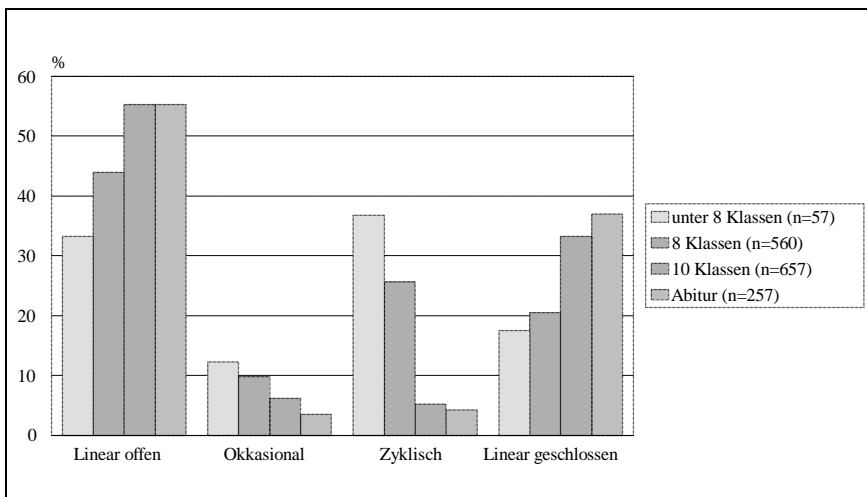


Abbildung 2: Zeitbewußtsein und Alter

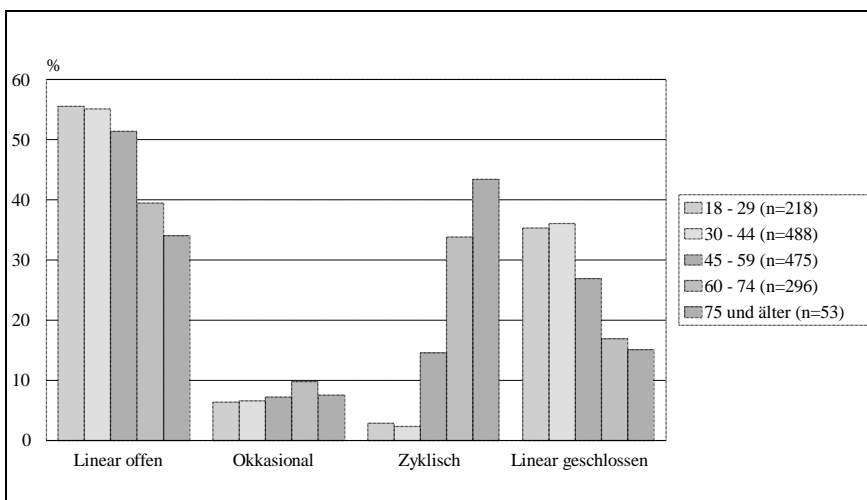
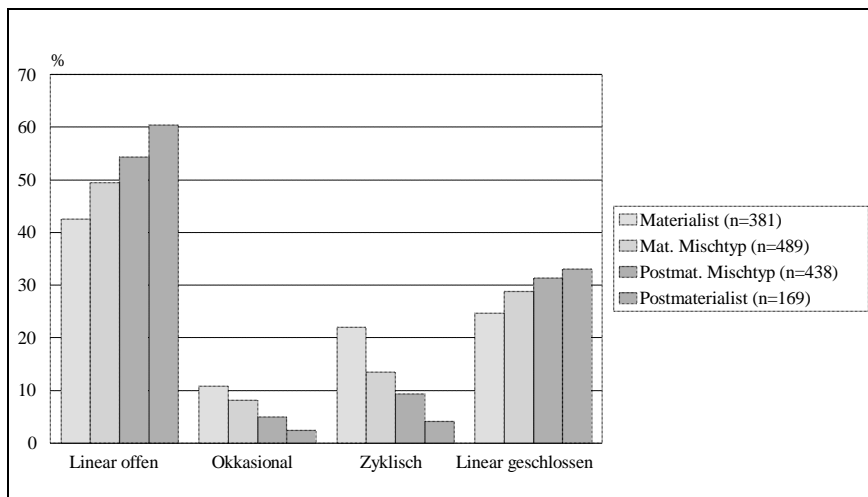


Abbildung 3: Zeitbewußtsein und Wertorientierung

Die Ergebnisse lassen sich folgendermaßen zusammenfassen:

1. Bildung hat einen deutlichen Einfluß auf das Zeitbewußtsein: Sowohl das okkasionale als auch das zyklische Zeitbewußtsein wird von Menschen mit niedrigem Bildungsstand (Abschluß unter oder 8 Klassen) vertreten. Das lineare Zeitbewußtsein nimmt dagegen mit zunehmender Bildung zu. Dabei zeigt sich, daß insbesondere das linear geschlossene Zeitbewußtsein eine Abhängigkeit mit Bildung aufweist. Der gleiche Zusammenhang besteht bei dem linear offenen Zeitbewußtsein, ist aber nicht so stark ausgeprägt.
2. Entsprechend sind die Zusammenhänge zwischen dem Zeitbewußtsein und Alter: Ein lineares Zeitbewußtsein wird vor allem von den jüngeren Personengruppen (18-29 Jahre, 30-44 Jahre) vertreten. Dabei zeigt sich ein leichter Unterschied hinsichtlich der Zukunftsorientierung: Während die Jungen (18-29 Jahre) noch eher unbestimmt in die Zukunft blicken, haben die Angehörigen der mittleren Alterskategorie (30-44 Jahre) ein konkretes Lebensziel vor Augen. Die älteren Personengruppen vertreten dagegen mehr ein okkasionales (60-74 Jahre) bzw. ein zyklisches Zeitbewußtsein (60-74 Jahre, 75 Jahre und älter).
3. Interessant ist der Zusammenhang zwischen dem Zeitbewußtsein und der Wertorientierung. Hier zeigte sich, daß Postmaterialisten überwiegend eine lineare

(sowohl offene als auch geschlossene) Zeitorientierung besitzen. Die Materialisten dagegen hängen eher einem okkasionalen und einem zyklischen Zeitverständnis an. Entsprechendes zeigt sich für die Mischtypen, wenn hier auch die Unterschiede nicht so deutlich sind: die materialistischen Mischtypen zeigen einen leichten Überhang bei dem okkasionalen Zeitbewußtsein; die postmaterialistischen Mischtypen hängen - wie die Postmaterialisten - eher dem linear offenen und dem linear geschlossenen Zeitbewußtsein an. Nun könnte man vermuten, daß dieser Zusammenhang sich aus der Alters- und Bildungsstruktur (s. die Punkte 1 und 2) ergibt, d.h. daß jüngere Menschen bzw. Menschen mit höherem Bildungsstand sowohl ein eher lineares Zeitverständnis haben als auch eher postmaterialistischen Werten anhängen. Dies erklärt aber nicht den hier gefundenen Zusammenhang: Auch wenn man Alter und Bildung bei der Kreuztabellierung zwischen Werten und Zeitbewußtsein kontrolliert, bleibt der Zusammenhang zwischen Zeitbewußtsein und Wertorientierung bestehen.

3. Worauf wirkt das Zeitbewußtsein?

Die Theorie geht davon aus, daß das Zeitbewußtsein - wie es hier verwendet wird - einen direkten Einfluß auf Handlungsorientierungen bzw. Handlungen ausübt. Es wird davon ausgegangen, daß insbesondere ein Zeitverständnis, das die Zukunft als wesentlichen Bestandteil berücksichtigt, bestimmte Handlungsimplicationen zur Folge hat. Hier ist ein Handeln zu erwarten, das auf die Zukunft gerichtet ist und aktiv gestaltend auf die Prozesse Einfluß nimmt. Es ist außerdem zu erwarten, daß dieses Zeitverständnis mit einem höheren Maß an Eigenverantwortung zusammenhängt.

Es wurden folgende Variablen als Indikatoren für Handlungen, Handlungsbereitschaft und wahrgenommene Eigenverantwortung gewählt:

1. Einstellung zur Pflegeversicherung. Diese Variable soll als Indikator für die wahrgenommene Eigenverantwortung verwendet werden. In dem Jahr, in dem die Erhebung der Studie stattfand, stand die Frage zur Debatte, wer für die Finanzierung der Pflegeversicherung zuständig sei. Diese Frage wurde auch im Fragebogen mit den folgenden Antwortmöglichkeiten gestellt: 1. „Vor allem sollte der Staat bezahlen.“ 2. „Jeder Einzelne sollte für sich selbst versorgen.“ 3. „Die Mittel sollten vor allem dadurch aufgebracht werden, daß alle etwas abgeben.“ 4. „Die Arbeitgeber sollten vor allem zahlen.“ Die Befragten mußten eine dieser Alternativen auswählen. Diese Variable kann insofern als Indikator für das Ausmaß an Eigenverantwortung gelten, als danach gefragt wird, ob man sich selbst (allein oder in der Gemeinschaft) oder aber andere (den Staat oder den Arbeitgeber) für die Finanzierung der

Pflegeversicherung zuständig sieht. Verantwortung kann individuell oder aber kollektiv getragen werden, d.h. die Verantwortung kann bei einem einzelnen Individuum bzw. Rollenträger oder bei einem Kollektiv liegen.³⁾ In diesem Sinne soll im folgenden nur die kollektive Handlungsverantwortung berücksichtigt werden, d.h. die Zuschreibung von Verantwortung an die Gemeinschaft oder aber an den Staat. Dieses Vorgehen erscheint vor allem deshalb angebracht, weil in der Befragung 85,4% der Fälle auf diese beiden Kategorien fielen. Man kann daraus schließen, daß im Fall der Pflegeversicherung die große Mehrheit der Befragten die Verantwortung für die Finanzierung als eine kollektive Verantwortung sieht - entweder liegt sie beim Staat oder aber bei der Gemeinschaft, der man selbst angehört.

2. Gegenwärtige Qualifizierung. Hier wurde danach gefragt, ob man sich gegenwärtig in irgendeiner Form von Qualifizierung befindet.⁴⁾ Aufgrund der geringen Fallzahl wird im folgenden nur unterschieden zwischen Personen, die sich auf irgendeine Art und Weise qualifizieren und Personen, die das nicht tun. Im Sinne der zeittheoretischen Überlegungen kann man davon ausgehen, daß insbesondere Qualifizierung eine Handlung darstellt, die ein Zukunftsbewußtsein reflektiert; man investiert jetzt in seine Ausbildung, um in der Zukunft den gewünschten Erfolg zu erreichen.
3. Bereitschaft zu sozialem Engagement. Eine weitere Frage in der Untersuchung beinhaltet die Bereitschaft zum sozialen Engagement. Hier wurde danach gefragt, ob man zur Mitarbeit in verschiedenen Organisationen bereit wäre, wobei eine Liste von möglichen Aufgaben vorgegeben wurde.⁵⁾ Eine vorgenommene Cluster-Analyse extrahierte hier drei verschiedene Ausprägungen: 1. Menschen, die bereit sind, sich für soziale, gemeinnützige Interessen wie z.B. für religiöse Ziele oder soziale Dienste einzusetzen (im folgenden als „soziale Interessen“ bezeichnet), 2. Menschen, die sich vor allem ihren eigenen Interessen widmen, wie z.B. politischen Ziele oder Freizeitinteressen (im folgenden als „utilitaristische Interessen“ bezeichnet) sowie 3. Menschen, die generell keine Neigung zeigen, sich auf diese Art zu engagieren.

Mit der Untersuchung dieser drei Variablen werden drei verschiedene Analyseebenen untersucht: Die Handlungsbereitschaft („Bereitschaft zum sozialen Engagement“), bereits erfolgte Handlungen („Qualifizierung“), sowie wahrgenommene Eigenverantwortung als motivationale Disposition („Finanzierung der Pflegeversicherung“). Handlungstheoretische Ansätze gehen davon aus, daß Persönlichkeitsmerkmale, wie sie hier durch die wahrgenommene Eigenverantwortung erfaßt wird, die motivationale Grundlage für Handeln bzw. Handlungsbereitschaft

bilden. Zu beachten ist daher, daß es sich bei den hier erfaßten Merkmalen um verschiedene Stufen im Handlungsprozeß handelt.

Zunächst sollen die Zusammenhänge zwischen Zeitbewußtsein und den hier aufgeführten Variablen jeweils getrennt analysiert werden. Dazu wurden insgesamt drei Kreuztabellierungen zwischen dem Zeitbewußtsein und der jeweiligen abhängigen Variablen vorgenommen. Die ermittelten Zusammenhänge sind sämtlich nach dem Chi-Quadrat-Test auf Unabhängigkeit auf dem 0,001-Niveau signifikant. Es zeigte sich folgendes Bild:

Tabelle1: Zusammenhänge von Eigenverantwortung, Qualifizierung und soziales Engagement mit Zeitbewußtsein

	Gesamt	Zeittypen			
		Linear offen	Okkasional	Zyklisch	Linear geschl.
<i>Eigenverantwortung</i>					
Staat	40,8	40,5	52,3	55,1	31,5
Allgemeinheit	59,2	59,5	47,7	44,9	68,5
χ^2 (3 df)	32,77				
N	1252	624	88	178	362
<i>Qualifizierung</i>					
nein	83,5	82,7	87,5	99,0	76,2
ja	16,5	17,3	12,5	1,0	23,8
χ^2 (3 df)	55,03				
N	1529	771	112	209	437
<i>Soziales Engagement</i>					
Utilitaristische Interessen	37,1	42,9	18,8	12,6	43,2
Soziale Interessen	18,0	17,6	21,4	15,0	19,4
nichts	44,9	39,5	59,8	72,4	37,4
χ^2 (6 df)	108,39				
N	1524	767	112	207	438

Angaben in Prozent (Spaltenprozent).

Diese Ergebnisse lassen sich folgendermaßen interpretieren:

1. Bezüglich der wahrgenommenen Eigenverantwortung wurde von der Vermutung ausgegangen, daß sowohl der linear geschlossene als auch der linear offene Zeittyp die Verantwortung für sich selbst übernehmen und entsprechend die Zuständigkeit für die Finanzierung der Pflegeversicherung bei der Gemeinschaft sehen. Von dem

zyklischen Zeittyp wird dagegen ein Abwälzen der Verantwortung auf den Staat erwartet. Diese Erwartungen bestätigen sich größtenteils. Der linear geschlossene Zeittyp sieht die Verantwortung insbesondere bei der Gemeinschaft, der zyklische und der okkasionale Zeittyp sehen dagegen den Staat als verantwortlich an. Über den linear offenen Zeittyp läßt sich keine Aussage treffen. Die Zusammenhänge bleiben stabil, wenn man mit Alter und Bildung kontrolliert.

2. Für die Qualifizierung ist zu vermuten, daß Personen mit linearem Zeitbewußtsein eher in die Zukunft planen und diese bei ihren jetzigen Handlungen mitberücksichtigen. Es ist zu erwarten, daß diese sich überdurchschnittlich qualifizieren, um so ihre beruflichen Aussichten für die Zukunft zu verbessern. Es zeigt sich hier, daß sowohl der linear offene als auch der linear geschlossene Zeittyp sich zu einem größeren Anteil in Qualifikation befindet als der zyklische und okkasionale Zeittyp. Der Zusammenhang zwischen dem linear-offenen Zeittyp und Qualifizierung geht allerdings verloren, wenn man Alter und Bildung kontrolliert. Da dieses Zeitbewußtsein vor allem in den jüngeren und gebildeteren Gruppen vorherrschend ist, ergibt sich zwangsläufig auch ein Zusammenhang mit Qualifizierung: Es sind vor allem jüngere Menschen bzw. Menschen mit einer höheren Bildung, die sich qualifizieren. Es scheint also so, daß ein lineares Zeitbewußtsein nur dann per se zu einem höheren Ausmaß an Qualifizierung führt, wenn man dabei auch ein konkretes Endziel vor Augen hat; diejenigen, die noch nicht wissen, wohin ihre Entwicklung sie führen wird, sehen nicht unbedingt die Notwendigkeit, in der Gegenwart durch Qualifizierung Vorsorge für später zu treffen. Dies ergibt sich eher aus ihrer gegenwärtigen Situation.
3. Auch bezüglich des sozialen Engagements besteht die Erwartung, daß Menschen mit linearem Zeitbewußtsein eine höhere Verantwortung empfinden und entsprechend eine höhere Bereitschaft zu sozialem Engagement aufweisen. Interessant ist, für welche Aufgaben diese Personen sich einsetzen würden: Die linear offenen Zeittypen sind lediglich bereit, sich für utilitaristische Interessen einzusetzen, während die linear geschlossenen Zeittypen sich neben diesen Interessen auch sozialen Aufgaben widmen würden. Dies entspricht der oben getroffenen Annahme, daß ein aktives Verhalten sich sowohl individuell, auf den eigenen Nutzen bezogen, als auch kollektiv, auf den Nutzen anderer bezogen, äußern kann. Da die Konzeption des linear geschlossenen Zeittyps davon ausgeht, daß möglicherweise ein religiöses Ziel als Endziel betrachtet wird, ist der Zusammenhang mit den sozialen Interessen durchaus plausibel. Zu vermuten ist auch, daß der linear offene Zeittyp, der mehr in der Gegenwart verwurzelt ist, eher an seinen eigenen Nutzen denkt und sich entsprechend engagiert. Die zyklischen als auch die okkasionalen Zeittypen dagegen würden sich generell eher nicht engagieren. Es besteht keine Tendenz zur Übernahme von sozialen oder utilitaristischen Aufgaben. Diese Zusammenhänge bleiben auch dann stabil, wenn man mit Bildung und Alter kontrolliert.

4. Zeitbewußtsein, Eigenverantwortung und Handlungsanstrengungen: eine multidimensionale Analyse

Es wurde gezeigt, daß die Form des Zeitbewußtseins einen signifikanten Effekt auf das Ausmaß an wahrgenommener Eigenverantwortung (hier erfaßt über die Einstellung zur Finanzierung der Pflegeversicherung) als auch auf bestimmte Handlungsdispositionen (Qualifizierung und Soziales Engagement) ausübt. Weiterhin wurde gezeigt, daß ein signifikanter Zusammenhang zwischen dem Zeitbewußtsein und bestimmten Wertorientierungen, hier erfaßt über den Inglehart-Index, bestehen. Lineare Zeittypen unterscheiden sich von den okkasionalen und zyklischen Zeittypen dadurch, daß sie eher postmaterialistische Werte vertreten, ein höheres Maß an Eigenverantwortung aufweisen und bezüglich ihrer Handlungsabsichten einen höheren Grad an Aktivität zeigen.

Im folgenden sollen die verschiedenen Zeittypen einer näheren Analyse unterworfen werden, indem sie anhand der hier erfaßten Merkmale - Wertedisposition, Eigenverantwortung und Handlungsbereitschaft - charakterisiert werden. Ziel ist dabei die Bildung von Gruppen, die sich bezüglich der Formen des Zeitbewußtseins möglichst voneinander unterscheiden. Dies soll zunächst mit einer explorativen Datenanalyse mit dem Programm SPSS CHAID geschehen, das Zusammenhänge zwischen den Variablen aufdeckt und anhand der untersuchten Merkmalen Untergruppen bildet. Anschließend sollen die so ermittelten Subgruppen mittels einer Konfigurationsfrequenzanalyse näher untersucht werden.

Das CHAID-Verfahren (Chi-squared Automatic Interaction Detector), das von Kass (1980) entwickelt wurde, ist ein exploratives Verfahren zur Untersuchung kategorialer Daten. Ausgangspunkt ist eine Menge von Merkmalsträgern, für die jeweils die Ausprägungen bestimmter kategorieller Merkmale Y , X_1 , ..., X_k erhoben wird, wobei Y als Zielvariable und X_1 , ..., X_k als erklärende Variablen dienen. Ziel des Verfahrens ist es, die Menge der Merkmalsträger mit Hilfe der erklärenden Variablen in Untergruppen (Segmente) aufzuteilen, die sich in Bezug auf die Zielvariable möglichst signifikant unterscheiden. Die Merkmalsträger werden dabei so zerlegt, daß es keinen Merkmalsträger gibt, der in zwei verschiedenen Segmenten oder aber in keinem Segment vorhanden ist (Magidson 1993: 3).

Die Segmentbildung basiert auf dem empirischen Signifikanzniveau eines Chi-Quadrat-Unabhängigkeitstests zwischen der Zielvariablen und den entsprechenden erklärenden Variablen. Die so entstehenden Segmente können damit als explizite Funktion der erklärenden Variable beschrieben werden (Magidson 1994: 125). Der CHAID-Algorithmus erfolgt in drei Phasen: der Merging-, Splitting- und Stopping-Phase. Dabei geht der

Algorithmus sequenzweise vor: Schrittweise werden neue Untergruppen gebildet, die dann wiederum weiter unterteilt werden. Jede der so gebildeten Untergruppen durchläuft jeweils die drei Phasen.

Merging-Phase: Für jede erklärende Variable werden die Kategorien so zusammengefaßt, daß diese sich möglichst signifikant in bezug auf die Zielvariable unterscheiden. Dies geschieht mittels des Chi-Quadrat-Tests auf Unabhängigkeit zwischen der Ziel- und der unabhängigen Variablen. Unterschreitet der p-Wert des Unabhängigkeitstests ein vorgegebenes Signifikanzniveau α_M (Merge-Level), so werden die entsprechenden Kategorien der betrachteten erklärenden Variable zusammengefaßt.

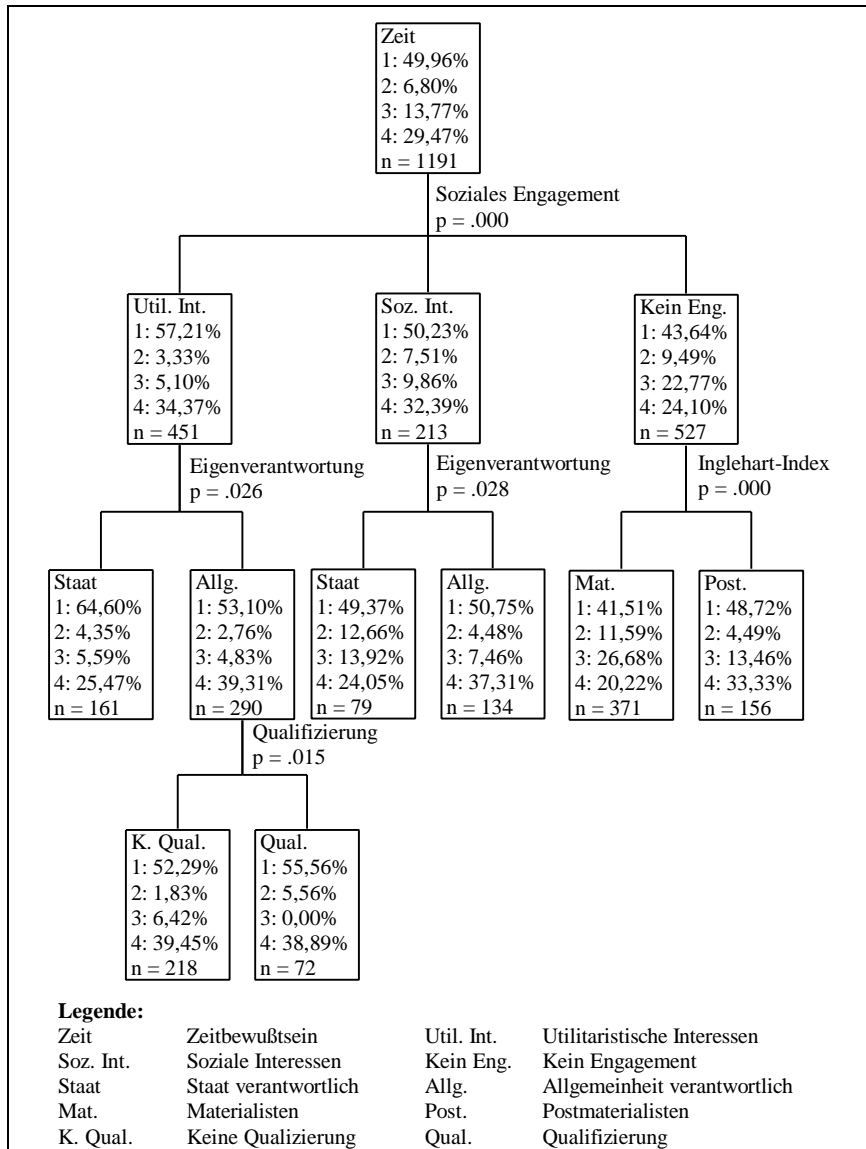
Splitting-Phase: Nachdem die Kategorien jeder erklärenden Variable zu neuen Kategorien zusammengefaßt wurden, wird nun die zu untersuchende Gruppe mit Hilfe der erklärenden Variable unterteilt, der bei einem Chi-Quadrat-Unabhängigkeitstest mit der Zielvariable zum kleinsten p-Wert führt. Eine Unterteilung der zu untersuchenden Gruppe erfolgt nur dann, wenn ein vorher festgelegtes Signifikanzniveau α_E (Eligibility-Level) unterschritten wird.

Stopping-Phase: Abschließend wird geprüft, ob die untersuchte Gruppe in der Splitting-Phase in Untergruppen unterteilt wurde. Sind diese Untergruppen gebildet, so wird für jede erneut die Merging-, Splitting- und Stopping-Phase durchlaufen. Die CHAID-Analyse ist dann beendet, wenn alle Untergruppen untersucht wurden. Zu kleine Untergruppen (CHAID verwendet hier als Vorgabe $n = 50$) werden allerdings aus dem Algorithmus ausgeschlossen, da die für den Unabhängigkeitstest verwendeten Prüfgrößen nur approximativ Chi-Quadrat-verteilt sind, und bei abnehmender Gruppengröße diese Approximation immer schlechter wird.

Im folgenden werden die Ergebnisse einer CHAID-Analyse dargestellt, die mit Zeitbewußtsein als Zielvariable und dem Inglehart-Index, Eigenverantwortung, Soziales Engagement und Qualifizierung als unabhängige Variable durchgeführt wurde. Es sei darauf hingewiesen, daß Zeitbewußtsein, wenn es hier als Zielvariable eingesetzt wird, nicht als zu erklärende Variable gilt. Die Chi-Quadrat-Tests gehen von einem nominalen Skalenniveau aus, es werden also lediglich *Zusammenhänge* erfaßt. Ziel ist somit nicht die Erklärung von Zeitbewußtsein, sondern lediglich die Bildung von Gruppen, die durch eine bestimmte Kombination von verschiedenen Merkmalausprägungen geprägt sind, wobei sich diese Gruppen hinsichtlich des Zeitbewußtseins möglichst voneinander unterscheiden. Die Darstellung erfolgt mittels eines Baum-Diagramms (Tree). Die Kästen stellen die Merkmalsträger (Knoten) dar, sie enthalten die prozentualen Anteile der jeweiligen Zeittypen sowie die Fallzahl der jeweiligen Gruppe. An den Zweigen steht jeweils die Variable, in der sich die Kategorien in bezug auf die abhängige

Variable am stärksten unterscheiden. Zusätzlich wird hier noch für jeden Split die Irrtumswahrscheinlichkeit p nach dem Chi-Quadrat-Unabhängigkeitstest angegeben.

Abbildung 4: Tree-Analysis



Zunächst einmal fällt auf, daß die erste Unterteilung anhand der Variable „Soziales Engagement“ vorgenommen wird; sie erbringt den größten Unterschied hinsichtlich der Häufigkeiten der vier verschiedenen Zeitbewußtseinsformen. Sowohl der linear offene als auch der linear geschlossene Zeittyp finden sich insbesondere bei den „Utilitaristischen Interessen“ wieder, der okkasionale und zyklische Zeittyp sind dagegen überdurchschnittlich in der Gruppe „kein Engagement“ vertreten. In einem nächsten Schritt werden die Segmente „Utilitaristische Interessen“ und „Soziale Interessen“ anhand der Variable „Eigenverantwortung“ weiter aufgesplittet. Für das Segment „kein Engagement“ dagegen erbringt das Einbeziehen des Inglehart-Indexes den größten Unterschied. Hier hat der CHAID-Algorithmus die Kategorien „Postmaterialisten“ und „Postmaterialistische Mischtypen“ sowie die Kategorien „Materialisten“ und „Materialistische Mischtypen“ zusammengefaßt. Eine weitere Aufspaltung wird lediglich in dem Segment „Utilitaristische Interessen-Allgemeinheit verantwortlich“ vorgenommen: hier wird noch die Qualifizierung als weitere bedeutsame unterscheidende Variable einbezogen. Es ist allerdings zu beachten, daß lediglich aufgrund der geringen Fallzahl in der Kategorie „Qualifizierung“ (N = 188, 15,6% der gesamten Fälle) diese Variable keine weitere Aufteilung ermöglicht. Setzt man die erforderliche Fallzahl der CHAID-Analyse herab, dann wird Qualifizierung auch in den anderen Segmenten miteinbezogen. Aus den oben genannten Gründen wird allerdings hier darauf verzichtet.

Wie lassen sich die gebildeten Segmente nun interpretieren? Es läßt sich erkennen, daß die verschiedenen Zeittypen in einigen Segmenten über-, in anderen dagegen unterrepräsentiert sind. Dabei sind außerdem die Unterschiede zwischen den Kategorien der jeweiligen Variablen zu beachten - sie ermöglichen einen Aufschluß darüber, welche Merkmalsausprägungen „typisch“ für den jeweiligen Zeittyp sind.

Entsprechend diesen „typischen“ Merkmalsausprägungskombinationen finden sich auch für die jeweiligen Zeittypen „atypische“ Kombinationen, also Segmente, in denen diese Zeittypen eher unterfrequentiert sind. Diese müssen nicht notwendig komplementär zu den typischen Kombinationen sein, wie man sehen kann. Für den zyklischen Zeittyp läßt sich beispielsweise eine besonders hohe Fallzahl in dem Segment „Kein Engagement-Materialistische Wertorientierung“ feststellen, in dem Segment „Utilitaristische Interessen-Allgemeinheit verantwortlich-Qualifizierung“ ist der zyklische Zeittyp dagegen überhaupt nicht vertreten. Ähnliche „typische“ bzw. „atypische“ Ausprägungen zeigen sich für die anderen Zeittypen.

Man erhält durch die Ergebnisse der CHAID-Analyse bereits einen ersten Eindruck davon, welche Zeittypen sich durch welche Kombinationen auszeichnen. Wie erwartet

zeichnen sich der zyklische und okkasionale Zeittyp insbesondere dadurch aus, daß sie sich nicht engagieren und zusätzlich eine eher materialistische Wertorientierung aufweisen. Die linearen Zeittypen dagegen engagieren sich sowohl für ihre eigenen Interessen als auch für das Wohl anderer, weisen zusätzlich eine eigenverantwortliche Einstellung auf und sorgen auch für die Zukunft, indem sie sich qualifizieren.

Im folgenden sollen die ermittelten Kombinationen der Merkmalsausprägungen näher untersucht werden. Das soll mit Hilfe der Durchführung einer Konfigurationsfrequenzanalyse (KFA), ein Verfahren, das von Lienert (1969) entwickelt wurde, geschehen (siehe auch Krauth/Lienert 1995). Bei der KFA handelt es sich um ein multivariates Analyseverfahren, mit dem Zusammenhänge zwischen kategorialen Daten ermittelt werden können. Ziel der KFA ist insbesondere das Ermitteln von sogenannten „Syndromen“. Hierbei handelt es sich um das Phänomen, daß zwei Variablen allein betrachtet keinen Zusammenhang aufweisen, bei Hinzunahme einer dritten oder sogar vierten Variable aber deutliche Interkorrelationen aufweisen. Es geht hier somit um das Auffinden von Tripel- oder Quadrupelassoziationen. Das Vorgehen der KFA entspricht dem CHAID-Verfahren insofern, als hier ebenfalls Individuen anhand mehrerer Merkmale in merkmalshomogene Subgruppen unterteilt werden. Die so entstehenden Systemkonfigurationen werden daraufhin untersucht, ob unter der Annahme der statistischen Unabhängigkeit der Variablen die beobachtete Zellhäufigkeit von der erwarteten abweicht. Die Prüfung wird mittels des Chi-Quadrat-Unabhängigkeitstestes vorgenommen. Dabei lassen sich sogenannte „Typen“ klassifizieren, d.h. Konfigurationen, die überfrequentiert sind, sowie sogenannte „Anti-Typen“, d.h. Konfigurationen, die unterfrequentiert sind. Es sei allerdings betont, daß dieses Verfahren keine mathematisch-statistische Absicherung der ermittelten Konfigurationen bietet; es ist lediglich ein heuristisches Hilfsmittel, mit dem Typen und Syndrome ermittelt werden können.

Es werden jeweils für die beiden in der CHAID-Analyse extrahierten Zweige („Zeit-Soziales Engagement-Eigenverantwortung-Qualifizierung“; „Zeit-Soziales Engagement-Inglehart-Index“) getrennte Konfigurationsfrequenzanalysen durchgeführt. Die nach dem Chi-Quadrat-Test signifikant über- bzw. unterbesetzten Zellen sind markiert, wobei die „Typen“ dunkel, die „Anti-Typen“ hell schraffiert dargestellt sind.

Bei der ersten durchgeführten KFA ist zunächst zu beachten, daß aufgrund der geringen Fallzahl in der Kategorie „Qualifizierung“ die erwartete Zellhäufigkeit den Wert 5 häufig unterschreitet. Von den 48 Zellen, die hier untersucht werden, liegt in 12 Zellen eine erwartete Häufigkeit unter 5 vor, das entspricht einem prozentualen Anteil von

25%. Von daher kann der gesamte Chi-Quadrat-Wert keine hinreichende Absicherung gegenüber der Nullhypothese leisten.

Tabelle 2: Konfigurationsfrequenzanalyse I

Zeityp	Soziales Engag.	Eigenverantwort.	Qualifiz.	Obs. Fr.	Exp. Fr.	P ² (1 df)
Linear-offen	Utilitaristisch	Staat	nein	84	77,19	0,60
Linear-offen	Utilitaristisch	Staat	ja	20	14,38	2,20
Linear-offen	Utilitaristisch	Allgemeinheit	nein	114	112,75	0,01
Linear-offen	Utilitaristisch	Allgemeinheit	ja	40	21,00	17,19***
Linear-offen	Sozial	Staat	nein	34	36,45	0,17
Linear-offen	Sozial	Staat	ja	5	6,79	0,47
Linear-offen	Sozial	Allgemeinheit	nein	60	53,25	0,86
Linear-offen	Sozial	Allgemeinheit	ja	8	9,92	0,37
Linear-offen	Nichts	Staat	nein	90	90,19	0,00
Linear-offen	Nichts	Staat	ja	8	16,80	4,61*
Linear-offen	Nichts	Allgemeinheit	nein	116	131,75	1,88
Linear-offen	Nichts	Allgemeinheit	ja	16	24,54	2,97
Okkasional	Utilitaristisch	Staat	nein	7	10,51	1,17
Okkasional	Utilitaristisch	Staat	ja	0	1,96	1,95
Okkasional	Utilitaristisch	Allgemeinheit	nein	4	15,35	8,39**
Okkasional	Utilitaristisch	Allgemeinheit	ja	4	2,86	0,46
Okkasional	Sozial	Staat	nein	10	4,96	5,11*
Okkasional	Sozial	Staat	ja	0	0,92	0,92
Okkasional	Sozial	Allgemeinheit	nein	4	7,25	1,46
Okkasional	Sozial	Allgemeinheit	ja	2	1,35	0,31
Okkasional	Nichts	Staat	nein	26	12,28	15,33***
Okkasional	Nichts	Staat	ja	1	2,29	0,72
Okkasional	Nichts	Allgemeinheit	nein	20	17,94	0,24
Okkasional	Nichts	Allgemeinheit	ja	3	3,34	0,03
Zyklisch	Utilitaristisch	Staat	nein	8	21,27	8,29**
Zyklisch	Utilitaristisch	Staat	ja	1	3,96	2,21
Zyklisch	Utilitaristisch	Allgemeinheit	nein	14	31,08	9,38**
Zyklisch	Utilitaristisch	Allgemeinheit	ja	0	5,79	5,78*
Zyklisch	Sozial	Staat	nein	11	10,05	0,09
Zyklisch	Sozial	Staat	ja	0	1,87	1,87
Zyklisch	Sozial	Allgemeinheit	nein	10	14,68	1,49
Zyklisch	Sozial	Allgemeinheit	ja	0	2,73	2,73
Zyklisch	Nichts	Staat	nein	68	24,86	74,86***
Zyklisch	Nichts	Staat	ja	0	4,63	4,63*
Zyklisch	Nichts	Allgemeinheit	nein	52	36,31	6,77**
Zyklisch	Nichts	Allgemeinheit	ja	0	6,76	6,76**
Linear-geschl.	Utilitaristisch	Staat	nein	28	45,53	6,75**
Linear-geschl.	Utilitaristisch	Staat	ja	13	8,48	2,41
Linear-geschl.	Utilitaristisch	Allgemeinheit	nein	86	66,51	5,71*
Linear-geschl.	Utilitaristisch	Allgemeinheit	ja	28	12,39	19,67***
Linear-geschl.	Sozial	Staat	nein	14	21,50	2,61
Linear-geschl.	Sozial	Staat	ja	5	4,01	0,24
Linear-geschl.	Sozial	Allgemeinheit	nein	37	31,41	0,99
Linear-geschl.	Sozial	Allgemeinheit	ja	13	5,85	8,74*
Linear-geschl.	Nichts	Staat	nein	44	53,21	1,59
Linear-geschl.	Nichts	Staat	ja	7	9,91	0,85
Linear-geschl.	Nichts	Allgemeinheit	nein	63	77,72	2,79
Linear-geschl.	Nichts	Allgemeinheit	ja	13	14,48	0,15

Gesamt Chi² = 244,85, df = 40, Signifikanzniveau: * = 0,05, ** = 0,01, *** = 0,001.

Tabelle 3: Konfigurationsfrequenzanalyse II

Zeittyp	Soziales Eng.	Inglehart-Index	Obs. Fr.	Exp. Fr.	P ² (1 df)
Linear-offen	Utilitaristisch	Materialist	115	132,61	2,34
Linear-offen	Utilitaristisch	Postmaterialist	143	92,70	27,30***
Linear-offen	Sozial	Materialist	56	62,63	0,70
Linear-offen	Sozial	Postmaterialist	51	43,78	1,19
Linear-offen	Nichts	Materialist	154	154,96	0,05
Linear-offen	Nichts	Postmaterialist	76	108,32	9,64**
Okkasional	Utilitaristisch	Materialist	9	18,05	4,54*
Okkasional	Utilitaristisch	Postmaterialist	6	12,62	3,47
Okkasional	Sozial	Materialist	9	8,53	0,02
Okkasional	Sozial	Postmaterialist	7	5,96	0,18
Okkasional	Nichts	Materialist	43	21,10	22,74***
Okkasional	Nichts	Postmaterialist	7	14,75	4,07*
Zyklisch	Utilitaristisch	Materialist	12	36,55	16,49***
Zyklisch	Utilitaristisch	Postmaterialist	11	25,55	8,29**
Zyklisch	Sozial	Materialist	15	17,26	0,30
Zyklisch	Sozial	Postmaterialist	6	12,07	3,05
Zyklisch	Nichts	Materialist	99	42,71	74,18***
Zyklisch	Nichts	Postmaterialist	21	29,86	2,63
Linear-gesch.	Utilitaristisch	Materialist	68	78,23	1,34
Linear-gesch.	Utilitaristisch	Postmaterialist	87	54,68	19,10***
Linear-gesch.	Sozial	Materialist	46	36,95	2,22
Linear-gesch.	Sozial	Postmaterialist	23	25,83	0,31
Linear-gesch.	Nichts	Materialist	75	91,41	2,95
Linear-gesch.	Nichts	Postmaterialist	52	63,90	2,21

Gesamt $\chi^2 = 209,27$, $df = 17$, Signifikanzniveau: * = 0.05, ** = 0.01, *** = 0.001.

Ein weiteres Problem, das generell bei der KFA zu berücksichtigen ist, besteht darin, daß hier simultan mehrere Tests am gleichen Material durchgeführt werden. Damit besteht eine Abhängigkeit der Tests untereinander, was zu einer faktischen Überschreitung der nominellen Irrtumswahrscheinlichkeit führt. Mit anderen Worten, die Wahrscheinlichkeit, die Nullhypothese abzulehnen, obwohl sie zutrifft, ist größer als der nominelle Alpha-Wert angibt.

Dieses Problem wird bei der simultanen Testdurchführung noch durch ein zweites Problem überlagert, das auch bei unabhängigen Tests auftritt. Führt man mehrere Tests mit jeweils derselben Nullhypothese bei einer festgelegten Irrtumswahrscheinlichkeit durch, so steigt die Wahrscheinlichkeit für mindestens ein signifikantes Ergebnis mit zunehmender Anzahl der Tests. Bei der simultanen Durchführung von beispielsweise

zehn unabhängigen Tests mit einer vorgegebenen Irrtumswahrscheinlichkeit $\alpha = 0,05$ wäre die Wahrscheinlichkeit für mindestens ein signifikantes Ergebnis bei jeweiliger Gültigkeit der Nullhypothese $= (1 - (1-\alpha)^{10}) = 0,401$. Bei einer inferenzstatistischen Sicherung von Typen würde dies dazu führen, daß mit hoher Wahrscheinlichkeit Typen nachgewiesen werden, auch wenn diese Inhomogenitäten sich zufallsmäßig erklären lassen würden. Um dieses Problem zu umgehen, soll hier ein Adjustierung des nominellen Alphawertes vorgenommen werden, so daß ein Typ bei Vorliegen der Nullhypothese nur mit der Wahrscheinlichkeit α^* auftritt. Es wird für alle r simultanen Tests die gleiche nominelle Irrtumswahrscheinlichkeit $\alpha_i = \alpha^*$ für $i = 1, \dots, r$ gesetzt. Dabei entscheidet man konservativ, indem man $\alpha^* = \alpha/r$ wählt. Somit würde man jeden einzelnen der r simultanen Tests nach der für α^* gültigen Signifikanzschranke beurteilen. Durch diese Alphaadjustierung wird sowohl der Effekt der Abhängigkeit von Tests als auch der Effekt der Häufung von Tests berücksichtigt. Für die hier vorgenommenen Konfigurationsfrequenzanalysen würde man also bei einer gewählten Irrtumswahrscheinlichkeit von 0,05 entsprechende Signifikanzschranken wählen: Für die KFA I ergibt sich $\alpha^* = 0,05/48 = 0,001$ und für die KFA II $\alpha^* = 0,05/24 = 0,002$.

Berücksichtigt man diese Alphaadjustierung, dann kann davon ausgegangen werden, daß von den oben festgestellten Typen und Anti-Typen nur die auf dem Signifikanzniveau von 0,001 auch tatsächlich statistisch bedeutsam sind. Demnach lassen sich in den beiden durchgeführten KFAs jeweils vier Typen diagnostizieren: Der linear offene Typ setzt sich für utilitaristische Interessen ein, hat eine eher eigenverantwortliche Persönlichkeit und qualifiziert sich. Hier scheint ein Syndrom vorzuliegen: Wie vorne festgestellt wurde, konnte hinsichtlich der wahrgenommenen Eigenverantwortung keine Aussage für den linear offenen Zeittyp getroffen werden. Hier zeigt sich nun - in der Kombination mit weiteren Merkmalen - daß dieser Zeittyp tatsächlich eher eine Eigenverantwortung wahrnimmt. Die gleiche Konstellation zeichnet sich für den linear geschlossenen Zeittypen ab: Er engagiert sich für utilitaristische Interessen, ist eher eigenverantwortlich und qualifiziert sich. Ein gegenläufiges Bild zeigt sich bei den zyklischen und okkasionalen Zeittypen: Beide sind dadurch charakterisiert, daß sie sich nicht engagieren, die Verantwortung beim Staat sehen und sich nicht qualifizieren. Auch bei der mit dem Inglehart-Index durchgeführten KFA zeichnen sich die erwarteten Typen ab: Die linearen Zeittypen setzen sich für utilitaristische Interessen ein und vertreten eher postmaterialistische Wertorientierungen. Die okkasionalen und zyklischen Zeittypen zeigen die erwartete Merkmalskombination von fehlender Engagementbereitschaft und materialistischen Wertorientierungen auf.

5. Zusammenfassung und Ausblick

Zusammenfassend kann festgestellt werden, daß entsprechend der getroffenen Hypothesen die vier verschiedenen Zeittypen sich hinsichtlich der untersuchten Merkmale - Wertorientierung, wahrgenommene Eigenverantwortung, Handlungsbereitschaft - signifikant voneinander unterscheiden. Die linearen Zeittypen zeichnen sich durch eine postmaterialistische Wertorientierung, wahrgenommene Eigenverantwortung und erhöhte Handlungsbereitschaft aus. Sowohl der okkasionale als auch der zyklische Zeittyp können durch eine materialistische Wertorientierung, fehlende Einsatzbereitschaft und Abwälzung der Verantwortung auf andere charakterisiert werden. Diese beiden Zeittypen unterscheiden sich somit nicht hinsichtlich der hier untersuchten Variablen, was verwundert. Von dem zyklischen Zeittyp wurden keine Handlungsanstrengungen erwartet, da er sich als fremdbestimmt empfindet. Der okkasionale Zeittyp dagegen ist nur nicht bereit, in die Zukunft zu investieren, da sein Lebensschwerpunkt in der Gegenwart liegt. Das schließt aber nicht aus, daß er sich für seine gegenwärtigen Interessen einsetzen würde - dies wurde hier aber widerlegt. Auch hinsichtlich der wahrgenommenen Eigenverantwortung spricht prinzipiell nichts dagegen, daß der okkasionale Zeittyp eine solche empfindet. Hier zeigte sich allerdings, daß er - ebenso wie der zyklische Zeittyp - die Verantwortung eher auf den Staat abschiebt.

Mit den hier verwendeten Variablen konnte eine charakterisierende Unterscheidung zwischen den linearen Zeittypen auf der einen Seite und den okkasionalen und zyklischen Zeittypen auf der anderen Seite erzielt werden. Es fehlen aber Merkmale, die dazu führen, daß auch zwischen dem linear offenen und linear geschlossenen sowie zwischen dem okkasionalen und zyklischen Zeittyp unterschieden werden kann. In dieser Untersuchung wurden insbesondere Variablen verwendet, die das Berufsleben betreffen. Als Indikator für die Eigenverantwortung wurde die Einstellung zur Finanzierung der Pflegeversicherung verwendet, ein Bereich, der als Bestandteil der Sozialen Sicherung auch mit dem Arbeitsbereich assoziiert wird. Die Handlungsanstrengungen wurden mittels der Qualifizierung und dem sozialen Engagement gemessen. Qualifizierung dient insbesondere der Verbesserung der Zukunftschancen für den beruflichen Werdegang. Das soziale Engagement, wie es hier erfaßt wurde, ist dadurch gekennzeichnet, daß eine Spaltung in soziale und eigennützige Interessen vorgenommen wurde; diese beinhalten wiederum auch berufs- und karrierebezogene Interessen. Es ist denkbar, daß die Verwendung von anderen Variablen, die beispielsweise Handlungsanstrengungen im familiären oder Freizeitbereich messen, hier eine deutlichere Unterscheidung zwischen dem okkasionalen und dem zyklischen Zeittyp erbringen würde. So könnte man vermuten, daß der

zyklische Zeittyp auch in diesen Bereichen eine fatalistische Haltung aufweist, der okkasionale Zeittyp sich dagegen eher engagieren würde. Hinsichtlich der linearen Zeittypen könnte man versuchen, Variablen zu finden, die stärker das Verfolgen von konkreten Zielen in der Zukunft hervorheben. Hier wäre zu erwarten, daß die linear geschlossenen Zeittypen bei Handlungen, die auf ein bestimmtes Ziel in der Zukunft gerichtet sind, eine höhere Handlungsanstrengung aufweisen.

In diesem Aufsatz wurde nur eine Perspektive von Handlungsbereitschaft dargestellt, es konnte aber doch gezeigt werden, daß Unterschiede zwischen den verschiedenen Zeitbewußtseinsformen vorliegen. Es wäre sicherlich interessant, auch andere Perspektiven zu untersuchen, um näheren Aufschluß über die Bedeutung des Zeitbewußtseins hinsichtlich der Handlungsbereitschaft zu erlangen.

Korrespondenzadresse

*Helga Sievers
Universität Leipzig
Institut für Soziologie
Augustusplatz 9
D - 04109 Leipzig*

Anmerkungen

*) Mein herzlicher Dank gilt Prof. Dr. Johannes Huinink und Dr. Michael Häder für ihre hilfreichen Anregungen und Kritik.

1) In der Erhebung „Leben Ostdeutschland“ wurden 1993 die vier verschiedenen Formen des Zeitverständnisses über folgende Items erhoben: 1. „Die Zukunft gestalte ich durch mein gegenwärtiges Verhalten. Ich weiß jedoch nicht genau, wie sie konkret aussehen wird.“ (Linear offen) 2. „Ich lebe vor allem im ‘Hier und Jetzt’. Bei dem was ich tue, interessiert mich kaum, was gestern war und was morgen passieren wird.“ (Okkasional) 3. „In meinem Leben ist eigentlich nichts Neues mehr zu erwarten. Da sowieso nur Sachen passieren, die es schon gegeben hat, ist es letztlich egal, wie ich mich verhalte.“ (Zyklisch) 4. „In meiner Entwicklung bewege ich mich auf ein bestimmtes, letztendliches Ziel zu, dessen Erreichen ich durch mein Handeln beeinflussen kann.“ (Linear geschlossen).

2) Es handelt sich bei der Untersuchung um eine Erhebungsreihe, die seit 1990 in Ostdeutschland mit einer jeweiligen Befragtenzahl von ungefähr 1500 Personen durchgeführt wurde. Diese wurde in Form einer Wiederholungsbefragung erhoben, in der mit Hilfe eines konstanten Frageprogramms der soziale Wandel in Ostdeutschland erforscht wird. Die Daten sind für die Jahre 1990 bis 1996 in Form eines kumulierten

Datensatzes beim Zentralarchiv Köln verfügbar (ZA-Nr. 6666), sowie für die jeweiligen einzelnen Erhebungen (für 1993: ZA-Nr. 6600). Codebücher liegen allerdings nur für den kumulierten Datensatz sowie für die Erhebung von 1990 vor.

3) Lenk (1992) geht von der Annahme aus, daß Verantwortung durch Zuschreibung entsteht; diese kann sowohl deskriptiver als auch normativer Natur sein. Er unterscheidet zwischen individueller und institutioneller Handlungsverantwortung, d.h. also Verantwortung wird entweder dem Einzelnen zugeschrieben oder aber irgendeiner Form von Gemeinschaft wie den Familien, Organisationen oder aber dem Staat.

4) Als mögliche Formen der Qualifizierung wurden hier genannt: „Studium an einer Hoch- bzw. Fachschule“, „Umschulung“, „Berufliche Fortbildung“, „Lehrgänge an Volkshochschulen“, „Kurse und Lehrgänge anderer Art“.

5) Hier wurden folgende Aufgaben in Organisationen aufgeführt: „Soziale, gemeinnützige Dienste“, „Politische Interessenvertretung“, „Freizeitinteressen“, „Vertretung von Arbeitnehmerinteressen“, „Religiöse Ziele“, „Spezielle Interessen von Gruppen“.

Literatur

Bierhoff, H. W., 1994: Verantwortung und altruistische Persönlichkeit. Zeitschrift für Sozialpsychologie 25: 217-226.

Coser, L. A./Coser, R. L., 1963: Time Perspective and Social Structure. S. 638-647 in: A. W. Gouldner/H. P. Gouldner (Hrsg.): Modern Sociology. New York.

Fischer, A., 1981: Zeitbewußtsein. S. 346-377 in: Shell AG, Jugendwerk, Jugend '81. Lebensläufe, Alltagskulturen, Zukunftsbilder. Opladen: Leske & Budrich.

Häder, M./Mohler, P. Ph., 1995: Zukunftsvorstellungen der Menschen als Erklärungsvariable für die Krise in der DDR und die gegenwärtige Situation in Ostdeutschland. Aus Politik und Zeitgeschichte. Beilage zur Wochenzeitung Das Parlament B27/95 vom 30. Juni 1995: 19-27.

Häder, M., 1996: Linear, Zyklisch oder Okkasional? Ein Indikator zur Ermittlung der individuell präferierten Form des Zeitbewußtseins. ZUMA-Nachrichten 39: 17-44.

Huinink, J./Häder, M., 1998: Der Einfluß des Alters und der Kohortenzugehörigkeit auf Kontrollverhalten, Zeitbewußtsein und Bedeutungszuweisung. S. 98-123 erscheint in: M. Häder/S. Häder: Sozialer Wandel in Ostdeutschland. Theoretische und methodische Beiträge zur Analyse der Situation seit 1990. Opladen: Westdeutscher Verlag.

Kass, G., 1980: An Exploratory Technique for Investigating Large Quantities of Categorical Data. Applied Statistics 29: 119 - 127.

Krauth, J./Lienert, G. A., 1995: Die Konfigurationsfrequenzanalyse (KFA) und ihre Anwendung in Psychologie und Medizin. Unveränderter Nachdruck der Ausgabe von 1973. Weinheim: Beltz.

Lenk, H., 1992: Zwischen Wissenschaft und Ethik. Frankfurt: Suhrkamp.

LeShans, L. L., 1952: Time Orientation and Social Class. *Journal of Abnormal and Social Psychology* 47: 589-592.

Lienert, G. 1969: Die „Konfigurationsfrequenzanalyse“ als Klassifikationsmittel in der klinischen Psychologie. S. 244-253 in: M. Irle (Hrsg.): Bericht 26. Kongreß der Deutschen Gesellschaft für Psychologie in Tübingen 1968. Göttingen: Hogrefe.

Luhmann, N., 1979: Zeit und Handlung - Eine vergessene Theorie. *Zeitschrift für Soziologie* 8: 63-81.

Magidson, J., 1993: SPSS for Windows CHAID Release 6.0. SPSS Inc. Chicago.

Rammstedt, O., 1975: Alltagsbewußtsein von Zeit. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 27: 47-63.

Wendorff, R., 1988: Konflikt und Koexistenz verschiedener Zeiten. S. 628-640 in: R. Zoll (Hrsg.): Zerstörung und Wiederaneignung von Zeit. Frankfurt/M.: Suhrkamp.

MITTEILUNGEN

TEXTPACK

Ein Programm zur computerunterstützten Text- und Inhaltsanalyse

Ab sofort steht eine Windows 95/NT Version des Programms TEXTPACK zur Verfügung. TEXTPACK ist ein Programmsystem zur computerunterstützten Text- und Inhaltsanalyse. Es ist sowohl für die quantitativ ausgerichtete Inhaltsanalyse als auch für das Datenmanagement innerhalb qualitativer Textanalysen geeignet. Der größte Teil der Prozeduren dient der Erschließung und Aufbereitung der Texte, zwei Prozeduren sind auf rein quantitatives Vorgehen ausgerichtet. Innerhalb des Bereichs der computerunterstützten Inhaltsanalyse (cui) ist TEXTPACK sowohl für die Analyse mit dem reinen Wörterbuchansatz (inhaltsanalytischer Diktionär) als auch mit dem empirischen Ansatz (z. B. Cluster der häufigsten Wörter) geeignet.

TEXTPACK erlaubt die Eingabe beliebiger ASCII-Zeichen im Textteil, die Verarbeitung von Stammwörtern (d.h. „endungslose“ Wörter) und Mehrwortverbindungen, die getrennte Auszählung der Worthäufigkeiten von Untereinheiten eines Textes in einem Arbeitsschritt, die Begrenzung der Ausgabedateien durch GO/STOP-Wortlisten und vieles andere mehr.

TEXTPACK bietet u.a. folgende Funktionen:

- *Worthäufigkeiten über den gesamten Text bzw. dessen Untereinheiten*

Die Datei der Worthäufigkeiten kann dadurch unterschiedlich sortiert und durch verschiedene Optionen in ihrem Umfang eingeschränkt werden.

- *Darstellung von Wörtern im Kontext*

Einzelne Wörter, Stammwörter oder Mehrwortverbindungen können in ihrem Kontext in der Form „Schlüsselwort im Kontext“ oder „Schlüsselwort außerhalb des Kontextes“ dargestellt werden.

- *Index der Wörter eines Textes*

- *Vergleich des Wortschatzes zweier Texte*
- *Codierung eines Textes mit Hilfe eines vorgegebenen Kategorienschemas (inhaltsanalytisches Diktionär)*

Es werden sowohl Dateien mit den Kategorienhäufigkeiten als auch mit den Kategorienabfolgen erzeugt. Die Validität der Vercodung kann mit mehreren Optionen überprüft werden.

- *Auswahl von Texteinheiten*

Es gibt zwei Möglichkeiten, auf einzelne Einheiten eines Textes zuzugreifen: über die Angabe der Identifikation oder über numerische Informationen in einer zweiten Datei, die als portable SPSS-Datei (EXPORT) oder eine Rohdaten-ASCII-Datei vorliegen können.

Neu gegenüber der DOS Version ist die Windows-Oberfläche von TEXTPACK. Die Funktionalität des Programms ist weitgehend gleich geblieben. TEXTPACK kann auf allen 486er PCs oder Pentium Rechnern installiert werden, die unter Windows 95 oder Windows NT laufen. Weitere Informationen sind bei Cornelia Züll (e-mail: textpack@zuma-mannheim.de) erhältlich.

Preise

Die Windows 95/NT Version von TEXTPACK kostet (in DM zuzgl. MwSt):

	nicht-kommerzielle Nutzung	kommerzielle Nutzung
Einzelkopie	200,-	1000,-
Institutskopie	1000,-	5000,-

SOZIALWISSENSCHAFTEN-BUS 1998/99

TERMINE UND PREISE

Erstmals im Jahre 1985 haben ZUMA und GFM-GETAS (Hamburg) den Sozialwissenschaften-Bus als sozialwissenschaftliche Mehrthemenumfrage eingerichtet. Er wird dreimal jährlich durchgeführt und ist zu einem festen Bestandteil der sozialwissenschaftlichen Umfrageforschung in Deutschland geworden. Möglich ist die Einschaltung von Einzelfragen bis hin zu Forschungsprojekten mit einer Dauer von circa 30 Minuten Einschaltzeit. Gemeinsam haben ZUMA und GFM-GETAS jetzt dieses Studienangebot auf die Möglichkeit zur Durchführung auch international vergleichender Studien erweitert. Damit soll deutschen Wissenschaftlern die Möglichkeit gegeben werden, bei Übernahme der Federführung und/oder Koordination von Studien mit ausländischen Partnern verstärkt auch die internationale Beratungsarbeit von ZUMA in Anspruch zu nehmen und das internationale Institutsnetz, in das GFM-GETAS als Mitglied der IPSOS-Institutsgruppe eingebunden ist zu nutzen. Der in der Tradition des ZUMA-BUS stehende Sozialwissenschaften-Bus hat folgende Besonderheiten gegenüber herkömmlichen Mehrthemenumfragen der Umfrageinstitute:

- Durch die Spezialisierung auf sozialwissenschaftliche Fragestellungen wird eine zu große Fragenheterogenität im Interviewablauf vermieden.
- Als Statistik wird das „Grundmodul“ der ZUMA-Standarddemographie erhoben. Dadurch ist eine den höheren Anforderungen der Sozialwissenschaften angemessene Auswertung möglich.
- Durch wahlweise Erweiterung der Standarddemographie um Spezialmodule kann der Forscher den Statistikteil um von ihm zusätzlich benötigte Demographieteile erweitern.
- Durch eine Recodierung auf die deutschen „Demographischen Standards“, zu denen die ZUMA-Standarddemographie kompatibel ist, ist der nationale Vergleich zu vielen Datensätzen der Sozial-, Media- und Marktforschung sowie zur amtlichen Statistik gegeben.
- Anders als bei den üblichen Buseinschaltungen findet vor Beginn des Hauptfeldes ein Pretest statt. Es werden jeweils 20 Interviews in West und Ost durchgeführt und ein Pretestbericht erstellt.
- Für die Durchführung der Feldarbeit gelten dieselben Erhebungsmodalitäten wie für methodisch anspruchsvolle sozialwissenschaftliche Exklusiv-Erhebungen.
- Durch Feldkontrollen und Datenbereinigung in einem bei Mehrthemenumfragen nicht branchenüblichen Ausmaß wird eine außergewöhnlich hohe Datenqualität gewährleistet.

- Der Sozialwissenschaften-Bus kann auch als Ausgangsstichprobe für Längsschnittstudien oder für vertiefende schriftliche oder telefonische Folgebefragungen genutzt werden, ebenso als Erhebungsinstrument für längerfristige Trendreihen.

Technische Daten

Der Sozialwissenschaften-Bus wird im Stichprobendesign des ADM-Master-Samples durchgeführt und sichert damit eine hohe regionale Repräsentanz über Bundesländer, politische Gemeindegrößenklassen und BIK-Regionsgrößenklassen.

Seit 1990, als der Sozialwissenschaften-Bus erstmalig auch in den neuen Bundesländern etabliert wurde, wird die Standard-Fallzahl auf insgesamt $n = 3.000$ Befragungen gestellt; davon entfallen in disproportionalen Ansatz $n = 2.000$ Fälle auf die westlichen und $n = 1.000$ Fälle auf die östlichen Bundesländer. Der Nachweis bundesweiter Gesamtergebnisse erfolgt dabei auf der Basis von $n = 2.500$.

Diese Anordnung hat viel Zuspruch gefunden, weil sie die Möglichkeit einschließt, auch in den neuen Bundesländern detaillierte Analysearbeiten durchzuführen, die bei einem proportionalem Stichprobenansatz (ca. 80% der Bevölkerung leben in den westlichen, ca. 20% der Bevölkerung in den östlichen Bundesländern) so nicht durchführbar sind. Diese „Standard-Belegung 1“ wird als größtes Stichprobenmodell für den Sozialwissenschaften-Bus angeboten.

Zusätzlich ist jetzt als verkleinertes Stichprobenmodell die „Standard-Belegung 2“ eingerichtet worden, die auf eine Gesamtfallzahl von $n = 2.000$ Befragungen gestellt wird. Die disproportionale Überhöhung der Fallzahl in den östlichen Bundesländern wird beibehalten; $n = 800$ Befragungsfälle entfallen auf das Gebiet „Ost“, $n = 1.200$ auf das Gebiet „West“; gesamtdeutsche Analysen können entsprechend auf eine Stichprobenbasis von $n = 1.500$ gestellt werden.

- **Grundgesamtheit:** In Privathaushalten lebende Bevölkerung im Alter ab 18 Jahren.
- **Stichprobengröße (Standard):** Zeitgleich in West und Ost durchgeführte Repräsentativ-Erhebungen mit 2.000 (1.200) Fällen West und 1.000 (800) Fällen Ost. Stichproben-Aufstockungen und Teilbelegungen möglich.
- **Stichprobe/Auswahlverfahren:** Je 1.000 Interviews: Einsatz eines ADM-Stichprobennetzes mit jeweils 210 Sampling-Points „West“ bzw. 192 Sampling-Points „Ost“. Auswahl der Befragungshaushalte nach strengem Random-Route, Auswahl der Befragungsperson im Haushalt durch Schwedenschlüssel-Verfahren.
- **Erhebungsmethode:** Mündliche Interviews.

- **Feldarbeit:** Pro Untersuchungsgebiet und entsprechend Stichprobengröße: Einsatz von 200 bis 400 Interviewern der GFM-GETAS, mit Erfahrung in der Durchführung sozialwissenschaftlicher Forschungsprojekte.
- **Ergebnisse:** Die Daten werden in Form eines bereinigten Datensatzes (iterativ gewichtet) auf Diskette entweder als ASCII - oder SPSS-System-Datei übergeben. Zusätzlich wird ein ausführlicher Methodenbericht über alle Einzelheiten der technischen Studiendurchführung erstellt.
- **Sonderwünsche/Modifikationen:** Auftragspezifische Sonderwünsche, wie z.B. Stichprobenerweiterungen (Regionale Aufstockung, Einbeziehung der ausländischen Bevölkerung, Herabsetzung der unteren Altersgrenze u.a.), oder eine Ergänzung des mündlichen Interviews mit schriftlichem Zusatzbogen oder die Übernahme von Auswertungsarbeiten (Grundauswertung und weiterführende Analyseverfahren, z.B. multivariate Auswertungen, sowie graphische Ergebnispräsentationen) und andere Spezifikationen können gegen gesonderte Rechnungstellung erfolgen.

Preise Sozialwissenschaften-Bus 1998/1999

Die Abrechnung für die einzelnen Einschaltungen erfolgt nicht nach Anzahl und Art der Einzel-Fragen, sondern nach dem insgesamt vorgesehenen resp. benötigten Befragungs-Zeitaufwand für den gesamten Einschaltkomplex. Der von uns zugrunde gelegte Minuten-Preis entspricht dabei im Mittel einem Fragenpreis von ca. 2.200 Mark in einer 2.000er Stichprobe. In diesem Preis ist bereits eine dem Gesamtaufwand der Einschaltung normalerweise entsprechende Anzahl von Befragungshilfen (z.B. Listen, Kärtchenspiele, Skalenvorlagen) eingeschlossen. Andere Stichprobenmodifikationen sowie die Einschaltung von Einzelfragen werden nach Aufwand kalkuliert. Bei Mehrfacheinschaltungen ist ein Sonderrabatt möglich.

Preise

Kosten pro Minute bei einem Einschalt-Umfang von	1998		1999	
	unter 15 Minuten	15 Minuten und mehr	unter 15 Minuten	15 Minuten und mehr
Standard-Belegung 1				
Gesamt-Bus, n = 3.000 (n = 2.000 West/n = 1.000 Ost)	6.950	6.600	7.100	6.750
Standard-Belegung 2				
Gesamt-Bus, n = 2.000 (n = 1.200 West/n = 800 Ost)	4.970	4.710	5.060	4.820
Sonderbelegungen				
n = 2.000 in nur einem Teilgebiet	5.220	4.950	5.310	5.100

n = 1.000 in nur einem Teilgebiet	2.670	2.530	2.720	2.630
Allen Preisen ist die gesetzliche Mehrwertsteuer hinzuzurechnen.				

Im Preis sind folgende Leistungen enthalten:

- Die Beratung bei der Fragenkonstruktion.
- Die Nutzung des Grundmoduls der ZUMA-Standarddemographie.
- Pretestdurchführung und -auswertung.
- Schreiben und Druck der Befragungsunterlagen (deutsche Version).
- Durchführung der Feldarbeit inklusive Interviewerkontrollen.
- Datenerfassung und -bereinigung.
- Übergabe der analysefähigen, iterativ gewichteten Datensätze wahlweise auf Magnetband oder auf Diskette entweder als ASCII- oder SPSS-System-Datei.
- Methodenbericht.
- Zusätzliche Designwünsche wie Splits, aufwendiges Vorlagematerial, postalische Zusatzbearbeitung etc., sowie die Texterfassung, das Kopieren oder die Vercodung von offenen Fragen oder Berufsangaben können gegen gesonderte Rechnungsstellung jederzeit eingerichtet werden.
- Für internationale Studien fordern Sie bitte gesonderte Kalkulationen an.

Termine

	Bus 3/98	Bus 1/99	Bus 2/99	Bus 3/99
Pretestbeginn	01.09.98	12.01.99	16.03.99	01.09.99
Hauptfeldbeginn	19.10.98	01.03.99	03.05.99	18.10.99
Hauptfeldende	09.12.98	17.04.99	19.06.99	08.12.99
Datenauslieferung	30.12.98	20.05.99	23.07.99	30.12.99

Der letzte Abgabetermin für die einzuschaltenden Fragen ist jeweils vier Wochen vor Pretest-Termin, bei Verzicht auf eine Pretest-Überprüfung 14 Tage vor Beginn des Hauptfeldes.

Zuständig für Anfragen und Beratung:

Dr. Jürgen H.P. Hoffmeyer-Zlotnik, ZUMA, Mannheim, Tel.: 0621-1246-175 (Fax: -100).

Sabine Haars, GFM-GETAS/WBA, Hamburg, Tel.: 040-80096-130 (Fax: -100).

BUCHBESPRECHUNGEN

Das Buch wendet sich sowohl an Sozialwissenschaftler, die Erfahrungen mit der quantitativen Textanalyse haben als auch an solche, die sich zum ersten Mal mit dem Thema befassen und eine Übersicht über diese Thematik suchen. Das Spektrum umfaßt u.a. Textanalyse-Anwendungen aus den Bereichen Kommunikationswissenschaften, Kultur-Analyse, psychologische Diagnose und Ereignisanalyse.

Carl W. Roberts (Hrsg.)

Text Analysis for the Social Sciences:

Methods for Drawing Statistical Inferences from Texts and Transcripts

Mahwah, New Jersey: Lawrence Erlbaum Associates 1997

ix+316 Seiten, ISBN 0-8058-1734-4, \$39,95 (paper)

Schwerpunkt des Buches sind Methoden und Anwendungen der quantitativen Textanalyse. Der Herausgeber definiert quantitative Textanalyse wie folgt: „If the method yields data matrices from which probabilistic inferences (i.e., P-values) can be legitimately drawn, the method is quantitative. Otherwise it is qualitative.“ (S. 2). Diese Definition wird von *G. Shapiro* und *J. Markoff* im ersten Kapitel verdeutlicht. Sie definieren die Inhaltsanalyse (Textanalyse) weiter als systematische Reduktion des Textes zu statistisch analysierbaren Symbolen, die das Vorkommen, die Intensität oder die Häufigkeit von sozialwissenschaftlich relevanten Merkmalen repräsentieren (S. 14).

Das Buch befaßt sich mit neuen Methoden-Entwicklungen der semantischen und der netzwerk-basierten Textanalyse und deren Kontrast zur traditionellen (themenorientierten) Textanalyse. Dazu werden Methoden und Anwendungsbeispiele, die in verschiedenen Zeitschriften veröffentlicht wurden, zusammengestellt und unter drei Methodenaspekten gruppiert (Themenanalyse, semantische und netzwerk-basierte Textanalyse).

Das Buch ist in drei Teile gegliedert: Methodologie, Anwendungsbeispiele und Ausblicke. Im ersten Teil werden die drei verschiedenen Methoden der Textanalyse ausführlich dargestellt. *Ph. J. Stone* gibt in „Thematic text analysis: new agendas for analyzing text content“ einen Überblick über die Entwicklung der Methode der computerunterstützten Themenanalyse. Themen sind definiert als sich wiederholende

Muster im Text, und die Themenanalyse identifiziert und codiert diese Muster. In seinem Beitrag geht *Stone* auf die Möglichkeiten dieses Verfahrens ein und beschreibt anhand von Beispielen aus der Marktforschung in den USA den Einsatz dieser Methode. Er verweist aber auch auf ihre Beschränkungen. In einem Ausblick zeigt er sich optimistisch über einen stärkeren Einsatz der computerunterstützten Textanalyse in Verbindung mit den Möglichkeiten des interaktiven Codierens. In einem zweiten Beitrag „Semantic text analysis: on the structure of linguistic ambiguity in ordinary discourse“ beschreibt *C. W. Roberts* die Methode der semantische Textanalyse. Im Gegensatz zur Themenanalyse wird hier die Relation zwischen Themen analysiert. Themen sind dabei Teile der semantisch definierten Grammatik, die auf Subjekt-Verb-Objekt Gruppen basieren. Zur Codierung gehören drei Stufen: die Stichprobenziehung, die Festlegung der Grammatik und die Codierung der Gruppen auf Basis dieser Grammatik. Dieses Vorgehen wird anhand eines Beispiels erläutert. Der dritte Beitrag des Methodenteils „Network text analysis: the network position of concepts“ von *K. M. Carley* beschreibt die netzwerk-basierte Textanalyse. *Carley* geht davon aus, daß Texte als Konzept-Netzwerke codiert werden können, wobei die Netzwerke kognitive Modelle repräsentieren. Daraus ergibt sich, daß ‘Bedeutungen’ codiert werden. Gesucht wird das regelmäßige Vorkommen von Bedeutungen (Konzepten) im Netzwerk. Der Beitrag endet mit einem detaillierten Beispiel eines Netzwerks und den dahinterstehenden Konzepten. Leider wird der Codierungsprozeß selbst nicht beschrieben.

Im zweiten Teil des Buches werden zu allen drei Methoden Anwendungsbeispiele gegeben. Ein Beispiel für eine Themenanalyse ist der Beitrag von *W. A. Danielson* und *D. L. Lasorsa* „Perceptions of social change: 100 years of front page content in the New York Times and the Los Angeles Times“. Mit der Analyse von Zeitungsartikeln, wie sie schon Lasswell in den 50er Jahren durchführte, wird sozialer und politischer Wandel in der amerikanischen Gesellschaft untersucht. Bei der Darstellung der Vorgehensweise fehlen genauere Informationen zur Idee des Samplings, was die Nachvollziehbarkeit der Analyse erschwert. In einem weiteren Kapitel geht *L. A. Gottschalk* in seinem Beitrag „The unobtrusive measurement of psychological states and traits“ auf die Gottschalk-Gleser Methode der Textanalyse ein.

Darauf folgen Beispiele zur semantischen Textanalyse. Den Anfang macht der Beitrag von *R. Franzosi* „Labor unrest in the Italian service sector: an application of semantic grammars“. Sein Ziel ist es, mit Hilfe einer semantischen Textanalyse italienischer Zeitungstexte, Arbeitskonflikte als ‘multiple-actor’ und ‘multiple-action’ Phänomene zu verstehen, die nicht mit Daten der offiziellen Streik-Statistik analysiert werden können. Der Schwerpunkt dieses Kapitels liegt in der Darstellung der Textanalyse-Ergebnisse.

S. Savaiano und *Ph. A. Schrodt* analysieren Nachrichtentexte aus dem Bürgerkrieg und der Hungersnot in Äthiopien. In ihrem Beitrag „Environmental change and conflict: analyzing the Ethiopian famine of 1984-1985“ zeigen sie, welche Verbindung es zwischen Umweltveränderungen und politischem Verhalten gibt. Dazu bestimmen sie Variablen, die interne Konflikte, Hungersnot und Migrationsverhalten messen. Mit Hilfe eines Parsers werden Ereignisse automatisch codiert, die die Variablen zur Messung des Konflikts bilden. Sie werden in der folgenden Analyse mit statistischen Daten zur Migration und zur Hungersnot kombiniert. In einem weiteren Kapitel von *E. M. Eltinge* „Assessing the portrayal of science as a process of inquiry in high school biology textbooks: an application of linguistic content analysis“ wird untersucht, ob und inwieweit in Biologie-Schulbüchern die Naturwissenschaften nicht länger als sterile Fakten vermittelt werden. Dieser Beitrag ist ein typisches Beispiel für eine semantische Textanalyse wie sie *Roberts* in seinem methodischen Beitrag in diesem Band beschreibt. Obwohl Ergebnisse der Analyse präsentiert werden, fehlt die Beschreibung der Konsequenzen, die sich aus der gewählten Textanalysemethode ergeben.

Als letzte Gruppe der Anwendungsbeispiele folgen Beiträge zur Netzwerk-Analyse. Im ersten Beitrag werden von *M. E. Palmquist*, *K. M. Carley* und *Th. A. Dale* in „Applications of computer-aided text analysis: analyzing literary and non-literary texts“ zwei Studien vorgestellt, die kognitive Modelle als Netzwerke darstellen. Themen der Studien sind zum einen das Bild von Robotern in der Science Fiction Literatur und zum anderen die Entwicklung einer gemeinsamen Wissensbasis eines Kurses. Obwohl in diesem Kapitel die Methode und die Ergebnisse ausführlich diskutiert werden, fehlt eine Diskussion der Komplexität der Netzwerke (Erstellung, Exploration). Der zweite Beitrag „Reasoning in economic discourse: an application of the network approach to the Dutch press“ von *J. Kleinnijenhuis*, *J. A. de Ridder* und *E. M. Rietberg* gibt ein Anwendungsbeispiel aus dem Bereich der Analyse von Wirtschaftsnachrichten in zwei niederländischen Zeitungen. Im Gegensatz zum vorherigen Beispiel, in dem kognitive Netzwerke vorgestellt werden, liegt hier der Schwerpunkt auf der Analyse von Ereignis-Netzwerken.

Den Abschluß dieses Teils des Buches bildet ein Kapitel von *R. Popping* zu Computerprogrammen zur Textanalyse. Popping selbst betont, daß die Softwareliste keinen vollständig Überblick über alle vorhandene Software gibt. Es wird versucht, alle aufgeführten Programme den drei Methoden zuzuordnen. Jedes Programm wird nur sehr kurz beschrieben.

Im dritten Teil des Buches wird ein Ausblick auf die Zukunft der Textanalyse versucht. Der erste Beitrag dieses Teils „The future of coders: human judgements in a world of

sophisticated software" von *G. Shapiro* diskutiert zwei unterschiedliche Analysemodelle: 'representational' und 'instrumental'. Dabei wird nach dem 'representational' Modell versucht herauszufinden, welche Inhalte der Autor/Sprecher gemeint hat. Dagegen versucht das zweite Modell ('instrumental'), die latente Bedeutung des Textes zu bestimmen. Der Beitrag diskutiert die Rolle der zukünftigen Analysesoftware und die Rolle der Codierer. Der Codierer wird dabei nicht als 'surrogate scientist', sondern als Instrument betrachtet, das Texte von einer natürlichen Sprache in eine künstliche Sprache (Kategorienschema) übersetzt. Im nächsten Kapitel von *R. Bechtel* „Developments in computer science with application to text analysis" wird ein Überblick über die Möglichkeiten der gegenwärtigen Computertechnologie für die Textanalyse gegeben. Vier Bereiche werden angesprochen: Verschriftung, Vorbereitung der Texte, Codieren und Analyse. Der Schwerpunkt des Artikels liegt vor allem in den ersten beiden Punkten. Leider wird der Teil Codieren und Analyse nur sehr knapp angesprochen und Hinweise z.B. zur Diktionär-Entwicklung fehlen. In „Text analysis and natural language database systems" betrachtet *A. L. Baker* Textanalyseysteme als Datenbanksysteme und beschreibt für jede oben genannte Methode ein Datenmodell und die damit verbundenen Abfragemöglichkeiten. Der Artikel befaßt sich ausschließlich mit Textmodellierung und -retrieval und geht nicht auf die Codierung ein.

Den Abschluß des Buches bildet ein Beitrag von *C. W. Roberts* „A theoretical map for selecting among text analysis methods", der die numerischen Ausgaben der drei Methoden der quantitativen Textanalyse beschreibt. Er versucht auf Grund der unterschiedlichen Ausgaben Empfehlungen zu geben, welche Methode für welche Forschungsfrage geeignet erscheint.

Das Buch ist für jeden, der sich einen Überblick über den aktuellen Stand der Methode der Textanalyse in den Sozialwissenschaften verschaffen möchte empfehlenswert, zumal derzeit praktisch kein vergleichbares Buch angeboten wird. Zudem gibt es einen Eindruck von der Vielfältigkeit der sozialwissenschaftlichen Textanalyse: es werden unterschiedliche Anwendungsbeispiele mit verschiedenen Textarten, verschiedenen Verfahren der Textauswahl, verschiedenen Codierebenen und Codierstrategien dargestellt. Es wäre allerdings wünschenswert, wenn im dritten Teil des Buches stärker die Codierungsfunktionalität von zukünftigen Systemen berücksichtigt würde. Die vom Herausgeber angestrebte klare Trennung der Methodenbeschreibungen und Anwendungsbeispiele ist hilfreich, aber es ist nicht in allen Beiträgen gelungen (wie z.B. im Methodenbeitrag von *Stone* und in den Anwendungsbeispielen von *Gottschalk* und *Popping*). Positiv sind der gemeinsame Index und die gemeinsame und detaillierte

Bibliographie für alle Beiträge. Alle Beiträge nehmen sowohl Bezug auf die methodischen Beiträge im ersten Teil als auch auf andere Beiträge im Buch.

Melina Alexa/Cornelia Züll

Das vorliegende Buch ist bereits die 4. Auflage der Einführung in SPSS der beiden Autoren. In den ZUMA-Nachrichten 37 (November 1995) wurde die zweite Auflage des Buches (damals zur SPSS Version 6.1) besprochen. Deshalb wird im folgenden nur eine kurze Zusammenfassung der damaligen Besprechung gegeben und vor allem auf die Neuerungen eingegangen.

Achim Bühl/Peter Zöfel

SPSS für Windows Version 7.5

Praxisorientierte Einführung in die moderne Datenanalyse

Addison-Wesley-Longman: Bonn 1998

669 Seiten, 79,90 Mark, ISBN 3-8273-1224-8

Die Autoren gehen zunächst kurz auf die historische Entwicklung von SPSS ein und beschreiben die Neuerungen der vorliegenden Version, bevor sie in einem einführenden Kapitel die Datenbeschreibung und -eingabe in SPSS vorstellen. Besonders ausführlich wird der Anwender mit der SPSS Umgebung und der Benutzerführung vertraut gemacht. Hier wird sehr detailliert auf die neuen Möglichkeiten der Verwaltung und Aufbereitung der SPSS-Ausgabe mit dem neuen Ausgabe-Navigator eingegangen: die Zugriffsmöglichkeiten auf einzelne Ausgabeteile sowie die Möglichkeiten, Tabellen zu verändern (Pivot-Technik) und zu formatieren. In einem späteren Kapitel wird gezeigt, wie die verschiedenen Ausgaben in Word für Windows übernommen oder in eine HTML-Datei umgesetzt werden können. In einem Anhang wird für alle SPSS Prozeduren zusammengestellt, ob diese schon Ausgaben im neuen Format des Navigators erstellen. Diese Kapitel des Buches beschreiben die wesentlichen Neuerungen für SPSS-Nutzer und sind sehr gut und ausführlich ausgefallen.

Nach der Erläuterung einiger statistischer Grundbegriffe werden einige der Verfahren zur Datenselektion und -modifikation vorgestellt. Danach werden die statistischen Verfahren an Beispielen demonstriert und die Ergebnisse interpretiert. Beispiele für eine grafische Umsetzung der Ergebnisse runden die Kapitel ab. Zum Buch gehört eine Diskette mit den Beispieldatensätzen, so daß alle Beispiele nachvollzogen werden können. Die statistischen Verfahren des Basis-Moduls werden ausführlich behandelt und die mathematisch-statistischen Grundlagen erläutert. Für die Prozeduren der Module „Advanced Statistics“ und „Professional Statistics“ sind die statistischen Erläuterungen recht knapp gehalten. Eine Liste der Literatur, in der die einzelnen statistischen Methoden beschrieben sind, wird im Anhang zur Verfügung gestellt.

Nach der Beschreibung der statistischen Verfahren wird in einem ausführlichen Kapitel auf die Weiterverarbeitungsmöglichkeiten der Grafik eingegangen. Nützlich sind vor allem die Hinweise, für welche Daten der jeweilige Grafiktyp sinnvoll einzusetzen ist.

Ein letztes Kapitel im Buch weist auf weitere Informationsquellen im Internet hin. Der Leser kann sich dort z.B. weitere Übungsaufgaben und Datendateien abrufen, oder er kann sich in einen Listserver zu SPSS-Fragen an der Universität Marburg eintragen.

Besonders gelungen ist den beiden Autoren die sehr ausführliche Beschreibung der Benutzeroberfläche von SPSS. Dabei wird auch das Arbeiten im Syntax-Fenster, das in anderen Beschreibungen häufig etwas zu kurz kommt, vorgestellt. Sehr ausführlich dargestellt wird die Weiterverarbeitung der Ausgabe im Navigator, die sich in SPSS 7.5 grundlegend geändert hat und die Übernahme der Ergebnisse in die Textverarbeitung. Deshalb eignet sich das Buch besonders für Leser, die den Umgang mit SPSS lernen wollen oder für solche, die eine ältere SPSS-Version kennen und nun zum ersten Mal SPSS 7.5 einsetzen wollen. Als reines Nachschlagewerk dagegen ist das Buch weniger geeignet, da nicht alle in SPSS zur Verfügung stehenden Optionen beschrieben werden. Auch der Statistikanfänger wird Schwierigkeiten haben, ohne begleitende Fachliteratur zurechtzukommen. Positiv ist, daß die Autoren auf jeweils neue Programmversionen von SPSS sehr schnell reagieren und schon kurz nach deren Erscheinen ein Buch zur aktuellen Version anbieten.

Cornelia Züll

Der von *Wolf Kirschner* verfaßte Beitrag 'Episentinel. Monitoring von Krankheiten und Gesundheitsrisiken in Deutschland auf der Grundlage einer Kreistypologie' ist ein Beispiel für eine empirisch anspruchsvolle Diskussion der Optimierung einer Stichprobenziehung im Themenfeld 'bevölkerungsbezogene Infektionsepidemiologie'. Diese Optimierung verfolgt einerseits den üblichen Zweck, durch ein Schichtungsmodell die Kosten-Ertragsrelation in bezug auf die untersuchte Stichprobe zu verbessern. Andererseits wird versucht, die Chance auf belastbare Aussagen der zu erwartenden Daten für reale geopolitische Einheiten unter der Bundesebene zu verbessern.

Der Beitrag liefert dabei zweierlei: (1) einen diskussionswürdigen Verfahrensvorschlag zur Entwicklung eines Schichtungsmodells und (2) ein nachvollziehbares konkretes Schichtungsmodell auf der Grundlage ausgewählter soziodemographischer bzw. sozioökonomischer Variablen (Ausgangsvariablenzahl: circa 100), das für den Bereich 'HIV/AIDS und STD's' unter der Maßgabe eines unveränderten Mengengerüsts in Deutschland quasi rezeptartig verwendet werden kann. Letzteres hat unseres Erachtens

für das angesprochene infektionsepidemiologische Teilgebiet einen unmittelbaren Wert. Der Verfahrensvorschlag zur Entwicklung eines Schichtungsmodells ist allerdings relativ ambitioniert und aufwendig. Es läßt sich sicherlich nicht in jedem Vorhabenskontext replizieren.

Wolf Kirschner

**Episentinel. Monitoring von Krankheiten und Gesundheitsrisiken in Deutschland
auf der Grundlage einer Kreistypologie**

München, Wien: Profil Verlag 1998

218 Seiten, ISBN 3-89019-440-0, das Buch kann über ZUMA kostenlos bezogen werden
(Kontakt: Jürgen H.P. Hoffmeyer-Zlotnik)

Unabhängig voneinander werden drei unterschiedliche Verfahren zur Bildung von Regionentypen eingesetzt, die sich u.a. zentral in der Art und Weise des Umgangs mit den Ausgangsvariablen unterscheiden (1: Transformation und Faktorenanalyse der Ausgangsvariablen/hierarchische Clusteranalyse, 2: qualitative Auswahl, Hauptkomponentenanalyse, 3: Variablenzahlreduktion durch 'Principal coordinates analysis/k-means Clusterverfahren). Die jeweiligen Ergebnisse werden gegenübergestellt, mit Blick auf 'Nutzungs- folgen' bei den infektionsepidemiologischen Zielvariablen bewertet und zu einer Schichtungsempfehlung verdichtet.

Da sich die 'Qualität' des konkret vorgeschlagenen Schichtungsmodells (fünf 'Schichten' in den alten Bundesländern, drei 'Schichten in den neuen Bundesländern') als stark zielvariablenabhängig darstellt und damit die vorgeschlagene konkrete Schichtung nur bedingt selbst für andere Infektionskrankheiten verwendet werden kann, relativiert sich die praktische Verwertbarkeit des Vorschlages.

Dies schmälert nicht seinen Wert für den wissenschaftlichen Diskurs zu einem zentralen Feld bevölkerungsbezogener Erhebungen: einer sachgerechten, effizienten wie effektiven Planung und Realisation der Untersuchungsgruppenauswahl.

Fritz Tiemann (Robert Koch-Institut, Berlin)

BUCHHINWEISE

BLICKPUNKT GESELLSCHAFT 4
SOZIALE UNGLEICHHEIT IN DEUTSCHLAND
Michael Braun und Peter Ph. Mohler (Hrsg.)
Opladen: Westdeutscher Verlag, 1998
315 Seiten, 59,80 Mark, ISBN 3-531-13200-8



Die im Band 4 der Reihe „Blickpunkt Gesellschaft“ vorgestellten Analysen beruhen zum größten Teil, aber nicht ausschließlich, auf der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS), die von ZUMA seit 1980 in West- und seit 1991 auch in Ostdeutschland durchgeführt wird. Die Beiträge decken ein weites Spektrum der Sozialberichterstattung und der Analyse des sozialen Wandels ab. Inhaltlich konzentrieren sie sich auf den Aspekt der sozialen Ungleichheit, insbesondere zwischen Ost- und Westdeutschland, der verschiedenen sozialen Schichten und zwischen den Geschlechtern. Die Autoren dieses Bandes sind *Stefan Weick, Peter H. Hartmann, Caroline Kramer, Michael Braun, Petra Hartmann, Karin Kurz, Jörg Blasius, Peter Ph. Mohler, Lorenz Gräf* und *Wolfgang Jagodzinski*.

ZUMA-NACHRICHTEN SPEZIAL BAND 3
CROSS-CULTURAL SURVEY EQUIVALENCE

Janet Harkness (Hrsg.)

Mannheim: ZUMA 1998

187 Seiten, 25 Mark, ISBN 3-924220-13-1

This volume, the third in the ZUMA-Nachrichten-Spezial series on methodological issues in empirical social science research, is devoted to issues of cross-cultural methodology. The focus is on issues of equivalence, the key requirement in cross-national and cross-cultural comparative research. As the contributions indicate, equivalence is, however, better thought of in terms of equivalencies - in social science surveys and in other standardised instruments of measurement. Contributors come from different countries and continents and from widely differing research backgrounds, ranging from linguistics to survey research and its methodologies, to cultural anthropology and cross-cultural psychology. They are, in order of appearance in the volume:



Timothy P. Johnson: Approaches to Equivalence in Cross-Cultural and Cross-National Survey Research. *Fons J.R. van de Vijver*: Towards a Theory of Bias and Equivalence. *Willem E. Saris*: The Effects of Measurement Error in Cross-Cultural Research. *Janet A. Harkness and Alicia Schoua-Glusberg*: Questionnaires in Translation. *Michael Braun and Jacqueline Scott*: Multidimensional Scaling and Equivalence: Is *having a job* the same as *working*? *Ingwer Borg*: A Facet-Theoretical Approach to Item Equivalency. *Peter Ph. Mohler, Tom W. Smith and Janet A. Harkness*: Respondents' Ratings of Expressions from Response Scales: A Two-Country, Two-Language Investigation on Equivalence and Translation.

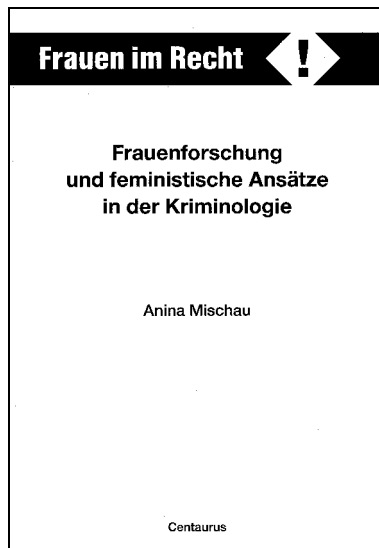
**FRAUENFORSCHUNG UND FEMINISTISCHE ANSÄTZE
IN DER KRIMINOLOGIE**

Dargestellt am Beispiel kriminologischer Theorien zur Kriminalität und
Kriminalisierung von Frauen

Anina Mischau

Centaurus 1997, 274 Seiten, 58 Mark, ISBN 3-8255-0150-7

Der weiblichen Kriminalität galt, aufgrund ihrer von der männlichen Kriminalität so verschiedenen quantitativen und qualitativen Struktur, von jeher das besondere Interesse der Kriminologen. Hiervon zeugt eine nahezu 100jährige Tradition kriminologischer Forschung und Theoriebildung zu diesem Thema. Die Autorin legt mit diesem Band erstmals eine umfassende theoriegeschichtliche Aufarbeitung deutschsprachiger Arbeiten zur Kriminalität und Kriminalisierung von Frauen vor. Bisherige Erklärungsmodelle, so die zentrale These der Autorin, beeindrucken eher durch ihre Quantität als durch ihre Qualität. Insbesondere „traditionelle“ Theorien verknüpfen geschlechtsspezifische Unterschiede bei kriminellem Verhalten und Kriminalisierungsprozessen eng mit der (männlichen) Ideologie von der Andersartigkeit der Frau und mit Mythen oder Halbwahrheiten über das biologische Geschlecht „weiblich“. Die Frauenforschung in der Kriminologie konnte in den letzten 20 Jahren wesentliche Widersprüche aufzeigen. Es ist ihr jedoch nicht gelungen, einen tatsächlichen Neuanatz zu formulieren. Die Autorin zeigt die Notwendigkeit einer neuen Auseinandersetzung mit diesen Theorien aus der Perspektive einer feministischen Wissenschaft(skritik) auf. Entscheidende Impulse hierfür geben neue Ansätze der feministischen Diskussion in der Kriminologie, die dezidiert nach den Möglichkeiten einer Integration der sozialen Kategorie Geschlecht in die kriminologische Theoriebildung fragen. Ziel der Autorin ist es, erste Gedanken für eine Kriminalitätstheorie zu skizzieren und so einer notwendigen Debatte neue Denkanstöße zu geben.



ZUMA-ARBEITSBERICHTE

Nachfolgend sind die ZUMA-Arbeitsberichte, die seit Dezember 1997 publiziert worden sind, in Form von Abstracts kurz dargestellt. ZUMA-Arbeitsberichte werden Interessenten auf Anfrage zugesandt. Bestellungen sind zu richten an:

Zentrum für Umfragen, Methoden und Analysen
ZUMA-Publikationen
Postfach 12 21 55
68072 Mannheim

* * *

Melina Alexa: Computer-assisted text analysis methodology in the social sciences. ZUMA-Arbeitsbericht 97/07.

The report presents an account of methods of research in computer-assisted text analysis in the social sciences. Rather than to provide a comprehensive enumeration of all computer-assisted text analysis investigations either directly or indirectly related to the social sciences using a quantitative and computer-assisted methodology as their text analytical tool, the aim of the report is to describe the current methodological standpoint of computer-assisted text analysis in the social sciences. The report provides, thus, a description and a discussion of the operations carried out in computer-assisted text analysis investigations. Both past and well-established as well as some of the current approaches in the field are examined and the techniques and the procedures involved are described. By this means, a first attempt is made toward cataloguing the kinds of supplementary information as well as computational support which are further required to expand the suitability and applicability of the method for the variety of text analysis goals.

* * *

Tracy L. Tuten: Getting a Foot in the Electronic Door: Understanding why People Read or Delete Electronic Mail. ZUMA-Arbeitsbericht 97/08.

For researchers, a key issue in developing electronic mail as a survey method is to understand what factors are at play as a potential respondent chooses to open or delete a mail message. This research investigated the process by which individuals make decisions about opening and reading versus deleting electronic mail and also assessed attitudes towards electronic mail surveys. The sample received an electronic mail message followed by a telephone interview. Results indicated that individuals delete

mail when the subject line does not interest them or when they do not recognize the name of the individual sending the email. Those interviewed reacted favorably to electronic mail surveys for scientific research despite issues of anonymity. However, respondents overwhelmingly described a dislike for commercially based electronic mail surveys.

* * *

Tracy L. Tuten: Electronic Methods of Collecting Survey Data: A Review of 'E-Research'. ZUMA-Arbeitsbericht 97/09.

Electronic data collection can take on many forms ranging from computer-administered surveys to internet surveys. While the benefits to collecting data electronically are many, obstacles also exist. This paper reviews three major methods of collecting survey data electronically: computer administered surveys, electronic mail surveys, and web surveys. The literature on these methods is reviewed, and the benefits and obstacles of each method are highlighted.

* * *

Heike Wirth und Paul Lüttinger: Die Klassenzugehörigkeit von Ehepaaren 1970 und 1993. Kontinuität oder Wandel? ZUMA-Arbeitsbericht 98/01. (Der Beitrag ist in überarbeiteter Form in der Kölner Zeitschrift für Soziologie und Sozialpsychologie erschienen: Heft 1, 1998, S. 47-77).

Die letzten Dekaden werden oft beschrieben als eine Phase der Auflösung traditioneller klassenspezifischer Strukturen und einer Zunahme individualisierter Lebensbeziehungen. Dies müßte sich auch als Abschwächung klassenspezifischer Heiratsbeziehungen bemerkbar machen. Für die Zeitpunkte 1970 und 1993 wurde diese Hypothese mit Daten der Volkszählung und des Mikrozensus überprüft. Es konnte eine weitgehende Stabilität in den Mustern der Klassenzugehörigkeit von verheirateten Männern und Frauen festgestellt werden. Zurückgeführt wird dies auf einen vor allem über soziale, kulturelle und regionale Faktoren stark selektierten und strukturierten Heiratsmarkt, der - entlang dieser Faktoren - in relativ homogene Subheiratsmärkte zerfällt, innerhalb derer die Partnerwahlen überwiegend stattfinden.

* * *

Rolf Porst: Im Vorfeld der Befragung. Planung, Fragebogenentwicklung, Pretesting. ZUMA-Arbeitsbericht 98/02.

Der Bericht gibt einen Überblick über die Planung und Durchführung sozialwissenschaftlicher Befragungen und die Konstruktion von Fragebogen. Darüber hinaus werden unterschiedliche Formen von Pretestverfahren beschrieben, die bei der Entwicklung eines Fragebogens zum Einsatz kommen können.

Rolf Porst: Erfahrung mit und Bewertung von Umfragen. Was unsere Befragten über Umfragen denken. ZUMA-Arbeitsbericht 98/03.

Im Rahmen einer bei ZUMA durchgeführten telefonischen Befragung wurden den Befragungsteilnehmern eine Reihe von Fragen gestellt, die sich auf ihre Erfahrung mit und ihre Bewertung von Umfragen bezogen. Dabei zeigte sich, daß mit 16 Prozent nur ein relativ geringer Anteil der Befragten bereits Erfahrung mit Befragungen hatte. Umfragen werden ganz allgemein als wichtig erachtet, als seriös und nach den Regeln des Datenschutzes durchgeführt; bezüglich der Seriosität und des Datenschutzes besteht bei teilnahmeerfahrenen Personen aber eine gewisse Skepsis. Die besondere Bedeutung des Interviewers für die Teilnahmeentscheidung wird erkennbar. Trotz einer positiven Grundhaltung der Befragten zu Umfragen wird aber auch die Befürchtung geäußert, aus Umfragen könne jeder genau das herauslesen, was ihm persönlich in den Kram passe.