

Programm Messung von Werten

Werte sind zentrale Konzepte der vergleichenden Kulturforschung und Gegenstand eines umfassenderen Forschungsprogramms. Das Unterprogramm *Entwicklung und Überprüfung von Instrumenten zur Messung von Werten* hat in der Methodenforschung der GESIS einen hohen Stellenwert. Nach einer weithin akzeptierten partiellen Charakterisierung sind Werte stabile, fest in die Persönlichkeit integrierte Maßstäbe oder Normen, die spezielle Einstellungen strukturieren und zumindest indirekt einen Einfluss unser Verhalten haben (Jagodzinski 2004). Aus diesen Merkmalen wird häufig abgeleitet, dass sich ein Wertewandel hauptsächlich als Generationenwandel und nicht als lebenszyklischer Wandel vollzieht. Spezifische Werte weisen weitere Eigenschaften auf: So konzipiert Inglehart Materialismus und Postmaterialismus als Pole eines eindimensionalen Kontinuums und Schwartz argumentiert, dass sich der Werteraum auf eine kleine Zahl von Dimensionen reduzieren lasse. Beide, Inglehart und Schwartz, gehen davon aus, dass die von ihnen postulierten Werte in allen Ländern und Kulturen nachweisbar sind.

Im Rahmen des Forschungsprogramms wird vorwiegend in Form von Sekundäranalysen überprüft, ob die in der Wertforschung etablierten Instrumente mit ausreichender Reliabilität latente Orientierungen messen, die die theoretisch geforderten Eigenschaften haben. Davidov (2008; Davidov et al. 2008; Davidov und Schmidt 2008) untersucht, ob die Werte von Schwartz mit den im ESS verwendeten Indikatoren in den beteiligten Ländern äquivalent gemessen werden kann. Markus Klein ist in Kooperation mit verschiedenen Kollegen der Frage nachgegangen, ob die von Inglehart favorisierten Forced-Choice Items tatsächlich weniger Messfehler produzieren als Rating-Skalen (vgl. zuletzt Klein et al. 2004a). Zweifel, ob der Postmaterialismusindex tatsächlich stabile Wertorientierungen misst, sind schon Ende der siebziger Jahre laut geworden, als sich zeigte, dass die Priorität des Ziels „Bekämpfung steigender Preise“ besonders empfindlich auf sich veränderte Inflationsraten reagiert (Böltken und Jagodzinski, 1985). Die Stabilität von Postmaterialismus und Postmaterialismus und die damit zusammenhängende Frage des Einflusses von Generationen-, Lebenszyklus- und Periodeneffekten wird später in Arbeiten von Jagodzinski (1984; 1996) und Klein (2003) untersucht.

Nachdem in den vergangenen Jahren die Bedeutung von Religion für die Entstehung von kulturellen Identitäten sichtbar geworden ist, wird momentan in diesem Programm die Äquivalenz von Skalen zur Messung von Religiosität untersucht.

Ausgewählte Publikationen:

Böltken, F./W. Jagodzinski, 1985: Insecure Value Orientations in an Environment of Insecurity. Postmaterialism in the European Community, 1970 to 1980. *Comparative Political Studies*, 17: 453-484.

Davidov, E., 2008: A Cross-Country and Cross-Time Comparison of the Human Values Measurements with the Second Round of the European Social Survey. *Survey Research Methods* 2(1), 33-46.

Davidov, E./B. Meuleman/J. Billiet/P. Schmidt, 2008: Values and Support for Immigration. A cross-country comparison (forthcoming). European Sociological Review (in press, published online).

Davidov, E./P. Schmidt, 2008: Are Values in the Benelux Countries Comparable? Testing for Equivalence with the European Social Survey 2004-5. In Loosveldt, G., M. Swyngedouw and Bart Cambré (Eds.), Measuring meaningful data in social research (Pp.373-386). Leuven: Acco.

Davidov, E./P. Schmidt/S. Schwartz. Bringing Values Back In: The Adequacy of the European Social Survey to Measure Values in 20 Countries. In press. Public Opinion Quarterly.

Jagodzinski, W., 1984: Wie transformiert man Labile In Stabile RELationen? Zur Persistenz postmaterialistischer Wertorientierungen. Zeitschrift für Soziologie, 13: 225-242.

Jagodzinski, W., 1996: The Metamorphosis of Life Cycle Change in Longitudinal Studies on Postmaterialism. S. 25-52. In: C. Hayashi/E. K. Scheuch (Hg.): Quantitative Social Research in Germany and Japan. Opladen: Leske + Budrich.

Jagodzinski, W., 2004: Methodological Problems of Value Research. Kapitel Fünf. S. 97-121. In: H. Vinken/ J. Soeters/P. Ester (Hg.): Comparing Cultures. Dimensions of Culture in a Comparative Perspective. Leiden: Brill.

Klein, M. und M. Pötschke, 2004 b): Die intraindividuelle Stabilität gesellschaftlicher Wertorientierungen – Eine Mehrebenenanalyse auf der Grundlage des sozio-ökonomischen Panels. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 56 (3): 432-456.

Klein, M., 2003: Gibt es die Generation Golf? Eine empirische Inspektion. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 55: 99-115.

Klein, M. /H. Dülmer/D. Ohr/M. Quandt/U. Rosar, 2004 a): Response sets in the Measurement of Values: a Comparison of Rating and Ranking Procedures. International Journal of Public Opinion Research, 16 (4).

© GESIS 01.09.2008

[http://www.gesis.org/forschung-lehre/programme-projekte/instrumente/messung-von-werten/?sword_list\[0\]=inglehart](http://www.gesis.org/forschung-lehre/programme-projekte/instrumente/messung-von-werten/?sword_list[0]=inglehart)

Inglehart Index

Value Orientation (Inglehart Items)

<http://www.gesis.org/dienstleistungen/daten/umfragedaten/eurobarometer/eb-trends-trend-files/list-of-trends/postmat/>

There is a lot of talk these days about what this country's goals should be for the next ten or fifteen years. On this card are listed some of the goals that different people say should be given top priority. Would you please say which of them you yourself consider to be most important in the long run ? And what would be your second choice?

Standard items (A):

1. **Maintaining order in the nation**
2. **Giving people more say in important government decisions**
3. **Fighting rising prices**

4. **Protecting freedom of speech**

Additional items (B):

1. **Maintaining a high level of economic growth**
2. **Making sure that this country has strong defence forces**
3. **Seeing that people have more say about how things are done at their jobs and in their communities**
4. **Trying to make our cities and countryside more beautiful**

Additional items (C):

1. **A stable economy**
2. **Progress towards a less impersonal and more humane society**
3. **Progress towards a society in which ideas count more than money**
4. **The fight against crime**

Additional items (D):

1. **Guarantee greater job security**
2. **Make our society more human**
3. **Raise salaries**
4. **Guarantee the participation of workers in the management of enterprises**

Usually the corresponding Materialism/Post-materialism (Inglehart) Index is included in the respective Eurobarometer data set. Please notice that the construction of this index is not consistent over time with respect to the treatment of MISSING VALUES (Don't Know / Not Ascertained to one or both basic variables). The original INRA data set editions include all MISSING VALUES into the index category MIXED. It is strongly recommended to check the codebook documentation and recode the missing values where necessary. For a complete discussion of this variable, see Ronald Inglehart: *The Silent Revolution: Changing Values and Political Styles among Western Publics*, Princeton University Press, 1977.

ZA Study Number	Eurobarometer	Fieldwork Month	Fieldwork Year	Question Number (1)	Variable Name (2)
0626	ECS70	2-3	1970	-	V63-V66(8)
0628	ECS73	9-10	1973	Q.33/Q.34	V27/V28 (3)
0989	5	5-6	1976	Q.235-Q.237	V111-V146 (4)
0990	6	11	1976	Q.117	V17/V18 (5)
0991	7	4-5	1977	Q.113	V9/V10 (5)
0992	8	10-11	1977	Q.114	V11/V12 (5)
0993	9	5-6	1978	Q.114	V11/V12 (5)
0994	10	10-11	1978	Q.119/Q.120	V16/V17 (5)
1036	11	4	1979	Q.157	V67/V70 (6)
1037	12	10	1979	Q.115	V12/V13 (7)
1038	13	4	1980	Q.135	V31/V32
1039	14	10-11	1980	Q.120	V14/V15
1206	15	4	1981	Q.121	V9/V10
1207	16	10-11	1981	Q.234	V49/V50
1208	17	3-4	1982	Q.146	V77/V78
1209	18	10	1982	Q.234	V35/V36
1318	19	3-4	1983	Q.247	V161/V162
1319	20	10	1983	Q.156	V38/V39
1320	21	3-4	1984	Q.139	V14/V15
1321	22	10-11	1984	Q.229	V127/V128
1541	23	3-4	1985	Q.135	V14/V15
1542	24	10-11	1985	Q.158	V20/V21

1543	25	3-4	1986	Q.126	V15/V16
1544	26	10-11	1986	Q.154	V45/V46
1712	27	4	1987	Q.135	V15/V16
1713	28	10-11	1987	Q.139	V21/V22
1714	29	3-4	1988	Q.138/Q.139	V67/V68 (3)
1715	30	10-11	1988	Q.128	V63/V64
1750	31	3-4	1989	Q.128	V15/V16
1751	31A	7	1989	Q.131	V20/V21
1752	32	10-11	1989	Q.7	V16/V17
1753	33	3-4	1990	Q.6	V16/V17
1960	34.0	10-11	1990	Q.14	V71/V72
1961	34.1	10-11	1990	Q.5	V20/V21
1962	34.2	10-11	1990	Q.6	V26/V27
2031	35.0	3	1991	Q.6	V22/V23
2032	35.1	3	1991	Q.4	V20/V21
2081	36	10-11	1991	Q.6	V23/V24
2141	37.0	3-4	1992	Q.6	V22/V23
2241	37.1	4-5	1992	Q.5	V21/V22
2242	37.2	4-5	1992	Q.3	V17/V18
2294	38.0	9-10	1992	Q.6	V22/V23
2346	39.0	3-4	1993	Q.6	V24/V25
2347	39.1	5-6	1993	Q.4	
2459	40	10-11	1993	Q.15	V32/V33
2563	42	11-12	1994	Q.3	V34/V35
2936	47.1	3-4	1997	Q.4	V39/V40
3205	52.1	11-12	1999	Q.12	

(1) With the exception of Eurobarometer 5 all question numbers and variables refer to the standard items (A).

(2) Please notice, that the variable names refer to the final ICPSR/ZA codebook editions, if available.

(3) all 12 items, asked separately in three groups (A/B/C) and all together including the least important choice

(4) all 12 items, multiple response format, asked for the most, second most and least important choice

(5) minor country specific variations in category 4

(6) standard items only asked in Denmark (V69/V70), all other countries (V67/V68) asked items (B)

(7) minor country specific variations in categories 2 and 4

(8) standard items with minor variations (V65/V66) and additional item group D (V63/V64)

© *GESIS* [Meinhard Moschner](#) 22.06.2001

Wertewandel/ Inglehart-Index

- Ronald Inglehart u.a.: World Values Survey Network: <http://wvs.isr.umich.edu/>
- Klein, Markus/Pötschke, Manuela (2000): Gibt es einen Wertewandel hin zum "reinen" Postmaterialismus? Eine Zeitreihenanalyse der Wertorientierungen der westdeutschen Bevölkerung zwischen 1970 und 1997. Zeitschrift für Soziologie, 29, S. 202-216

- Bauer-Kaase, P., & Kaase, M. (1998). Werte und Wertewandel - ein altes Thema und eine neue Facette. In *Empirische Forschung und wirtschaftspolitische Beratung. Festschrift für Hans-Jürgen Krupp zum 65. Geburtstag (256-274)*. Frankfurt: Campus.
- Inglehart, R. (1993). *The Silent Revolution in Europe - intergenerational change in post-industrial societies*. New York:.
- Klages, H., & Gensicke, T. (1999). *Wertewandel und bürgerschaftliches Engagement an der Schwelle zum 21. Jahrhundert*. Speyer: Forschungsinst. für Öffentl. Verwaltung.
- Klein, M. A., Kai. (1999). Ranking- und Rating-Verfahren zur Messung von Wertorientierungen, untersucht am Beispiel des Inglehart-Index. Empirische Befunde eines Methodenexperiments. In. <http://www.politik.un-mainz.de/kai.arzheimer/inglehart-panel>..
- Klein, M., & Pötschke, M. (2000). Gibt es einen Wertewandel hin zum "reinen" Postmaterialismus? Eine Zeitreihenanalyse der Wertorientierungen der westdeutschen Bevölkerung zwischen 1970 und 1997. In *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (202-216). Köln: Opladen.
- Witte, E. H. (1996). Wertewandel in der Bundesrepublik Deutschland (West) zwischen 1973 und 1992. Alternative Interpretationen zum Inglehart-Index. In *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (534-541). Köln: Opladen.

TITEL: Ranking- und Rating-Verfahren zur Messung von Wertorientierungen, untersucht am Beispiel des Inglehart-Index. Empirische Befunde eines Methodenexperiments

ENGL. TITEL: Ranking and Rating Procedures for the Measurement of Values, Analyzed with an Example of the Inglehart-Index. Empirical Results of a Methodological Experiment .

QUELLE: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 51, 1999, S. 550-564

VOLLTEXT

ZUSAMMENFASSUNG: Über die Frage, ob Ranking-Prozeduren wirklich besser für die Messung von Wertorientierungen geeignet sind als Rating-Skalen, gibt es in der empirischen Sozialforschung eine andauernde Diskussion. Um diese Debatte mit mehr empirischer Evidenz anzureichern, haben wir ein Methodenexperiment in eine zweiwellige schriftliche Panelbefragung integriert: In der ersten Panelwelle wurden die Befragten gebeten, die vier Items des Inglehart-Index sowohl unabhängig voneinander in ihrer Wichtigkeit zu bewerten als auch rangzuordnen. In der zweiten Panelwelle haben wir diese Bitte wiederholt, dabei allerdings der Hälfte der Respondenten einen Fragebogen vorgelegt, in dem die Reihenfolge der Items vertauscht war. Mit Hilfe dieses experimentellen Designs können wir zeigen, dass ein Zusammenhang zwischen der Reihenfolge der Items im Fragebogen und den von den Befragten vergebenen Rangpositionen besteht. Dies gilt in besonderem Maße für Befragte, die in der ersten Panelwelle im Rahmen des Rating-Verfahrens allen Items dieselbe Wichtigkeit zugeschrieben hatten. Das Ranking-Verfahren ist also für Response-Sets ebenso anfällig wie Rating-Skalen. Vor dem Hintergrund dieser Befunde besteht kein Grund für die Annahme, dass das Ranking dem Rating in methodischer Hinsicht überlegen sei.

ABSTRACT: In empirical social research there is a long debate about the question, whether ranking procedures are indeed more appropriate for the measurement of value orientations than rating scales. To gain empirical clarity, we incorporated a methodological experiment into a two-wave mail panel survey. In the first wave we asked respondents to rank and rate the four items of the well known Inglehart-Index. In the second wave we split the sample and while asking the same question, we changed the order of the items in the questionnaire for half of the respondents. By means of this experimental design we can demonstrate that there exists a correlation between the

order of the items in the questionnaire and the rank-order that respondents assign to these items. This relationship is even stronger for respondents which gave the same rating to all items in the first panel wave. Thus, ranking procedures seem to be affected by response sets as well as rating procedures. In the light of these findings, there remains no basis for the assumption, that rankings are superior to ratings in methodological terms.

KORRESPONDENZANSCHRIFT: Dipl.-Volksw. Markus Klein M.A., Zentralarchiv für empirische Sozialforschung an der Universität zu Köln, Bachemerstr. 40, D-50931 Köln, E-Mail: klein@za.uni-koeln.de; [Kai Arzheimer](#), Institut für Politikwissenschaft, Johannes Gutenberg-Universität, Colonel-Kleinmann-Weg 2, D-55099 Mainz

<http://www.politik.uni-mainz.de/kai.arzheimer/ingelhart-panel/ingelhart-panel.html>

<http://www.kai-arzheimer.com/ingelhart-panel/ingelhart-panel.html> (am 27_02_09)

jetzt hier – University of Essex / sowie eigene Homepage -

<http://www.essex.ac.uk/government/staff/academic/arzheimerK.shtm>

<http://www.kai-arzheimer.com/Welcome.html>

Ranking- und Rating-Verfahren zur Messung von Wertorientierungen, untersucht am Beispiel des Inglehart-Index
Empirische Befunde eines Methodenexperiments

Dipl.-Volkswirt Markus Klein M.A.

Zentralarchiv für Empirische

Sozialforschung der Universität zu Köln

Postfach 41 09 60

50 869 Köln

Tel.: 0221 / 470-3157

klein@za.uni-koeln.de

[Kai Arzheimer](#)

Institut für Politikwissenschaft

Abt. Innenpolitik und politische Soziologie

Johannes Gutenberg-Universität

55 099 Mainz

Tel.: 06131 / 39-3782

Ranking- und Rating-Verfahren zur Messung von Wertorientierungen, untersucht am Beispiel des Inglehart-Index*

Empirische Befunde eines Methodenexperiments

Markus Klein / Kai Arzheimer

1. Einleitung und Problemstellung

Welches Abfrageformat für die Erhebung von Wertorientierungen angemessen ist, wird in der empirischen Sozialforschung seit jeher kontrovers diskutiert (vgl. u.a. Hicks 1970; Rokeach 1973; Munson/McIntyre 1979; Rankin/Grube 1980; Reynolds/Jolly 1980; van Deth 1983; Alwin/Krosnick 1985; Krosnick/Alwin 1988; Maag 1991; Bean/Papadikis 1994a,b; Bürklin/Klein/Ruß 1994, 1996; Hellevik 1994; Inglehart 1994; Russel/Gray 1994; Klein 1995; Inglehart/Klingemann 1996). Im Mittelpunkt steht dabei die Frage, ob die Indikatoren zur Erfassung der verschiedenen Werte von den Befragten einzeln in ihrer Wichtigkeit beurteilt werden sollen (*Rating-Verfahren*) oder ob es sinnvoller ist, die Wertindikatoren nach Maßgabe ihrer Wünschbarkeit rangordnen zu lassen (*Ranking-Verfahren*). In der Praxis der Umfrageforschung hat sich im wesentlichen das Ranking-Verfahren durchgesetzt. Dieses liegt

auch dem wichtigsten Instrument zur Erfassung von Wertorientierungen zugrunde, dem sog. *Inglehart-Index*.

Bitte beachten Sie: Es handelt sich bei diesem Text nicht um die endgültige Druckfassung, sondern um ein Manuskript. Bitte zitieren Sie deshalb nur nach der gedruckten Fassung!

Im vorliegenden Beitrag werden wir die Ergebnisse eines Methodenexperiments berichten, in dessen Rahmen der Inglehart-Index sowohl über das Ranking- als auch über das Rating-Verfahren erhoben wurde. Für die 12-Item-Version des Inglehart-Index ist dies zwar bereits von anderen Autoren geleistet worden (van Deth 1983; Bean/Papadikis 1994), wir haben dieses Experiment im Rahmen einer zweiwelligen Panelbefragung nun aber auch für die in der Forschungspraxis sehr viel bedeutsamere 4-Item-Version wiederholt. In der zweiten Panelwelle haben wir darüber hinaus im Rahmen eines Split-half-Experiments die Antwortvorgaben rotiert, um empirisch prüfen zu können, ob sich die dem Ranking-Verfahren zugeschriebene Eigenschaft, reflektiertere Urteile der Befragungspersonen zu erzwingen, auch empirisch nachweisen läßt. Zunächst aber wollen wir kurz diskutieren, welche methodischen Vor- und Nachteile den beiden Erhebungsverfahren in der Literatur zugeschrieben werden und welche theoretischen Vorentscheidungen über die Struktur und die Funktion gesellschaftlicher Wertorientierungen ihnen jeweils zugrunde liegen.

2. Ranking- und Rating-Verfahren zur Messung von Wertorientierungen

2.1 Das Ranking-Verfahren

Der wichtigste *methodische Vorteil* des Ranking-Verfahrens wird darin gesehen, daß es sich nicht leicht "abarbeiten" läßt. "Using rankings instead of ratings prevents the respondents from taking the easy way of scoring everything as important" (van Deth 1983: 410). Aufgrund des "Forced-choice"-Formates müßten sich die Befragten vielmehr Gedanken über die relativen Wichtigkeiten der verschiedenen Werte machen. Das Ranking-Verfahren verlange von den Befragten also ein Mindestmaß an kognitivem Aufwand, der sich in reflektierteren Urteilen der Befragungspersonen niederschlägt und somit zu einer besseren Datenqualität führe (Krosnick/Alwin 1988: 534; Alwin/Krosnick 1985: 537). Diese zeige sich u.a. in der für Ranking-Techniken geringfügig höheren Test-Retest-Stabilität (Munson/McIntyre 1979; Rankin/Grube 1980).

Als *methodischer Nachteil* des Ranking-Verfahrens wird angeführt, daß es den Befragten eine relativ komplizierte Aufgabe abverlangt, die unter Umständen eine Überforderung bedeute (Krosnick/Alwin 1988: 527). Ranking-Prozeduren sind darüber hinaus in der Erhebung zeitaufwendig und insofern umfragetechnisch eine relativ "teure" Lösung (Alwin/Krosnick 1985: 536). Als weiterer Nachteil ist anzusehen, daß Ranking-Verfahren im Rahmen von Telefoninterviews nur schwer umzusetzen sind, da die Rangordnungsaufgabe oftmals nur mit Hilfe visueller Unterstützung durch ein Kartenspiel möglich ist (Alwin/Krosnick 1985: 536).

Ein sehr viel grundlegenderer Einwand gegen das Ranking-Verfahren besteht darin, daß die Rangordnung von Items nur für das einzelne Individuum Bedeutung besitzt, nicht aber für den interindividuellen Vergleich (Hicks 1970; van Deth 1983: 411). Außerdem ist die Analyse von Ranking-Daten mit statistischen Problemen behaftet: Da die Summe der vergebenen Rangpositionen für jeden Befragten eine Konstante ergibt, handelt es sich beim Ranking um eine sog. *ipsative* Messung, d.h. die Meßwerte für die einzelnen Wertitems sind nicht unabhängig voneinander. Ist erst einmal ein Item auf den ersten Rangplatz gesetzt, können die anderen Items nur noch auf den verbleibenden niedrigeren Rangplätzen eingestuft werden, da Rangplatzbindungen (ties) nicht möglich sind. Dies bedeutet, daß zwischen den Rangpositionen der Items notwendigerweise negative Korrelationen auftreten müssen (Hicks 1970;

Alwin/Krosnick 1985: 536; Krosnick/Alwin 1988: 528; Munson/McIntyre 1979: 49; van Deth 1983: 408; Rankin/Grube 1980: 234, 244). Diese Eigenschaft der Ranking-Skalen führt dazu, daß statistische Verfahren, die auf der Analyse von Kovarianzstrukturen basieren, nicht angewendet werden können. "In fact, hardly any statistical technique can be applied legitimately in these cases" (van Deth 1983: 411). Die Ipsativität der Messung führt außerdem zu Problemen bei der Bestimmung der Dimensionalität des Werteraumes, da sie die Differenzierung der Items entlang einer einzigen bipolaren Dimension forciert (Flanagan 1987).

Die *theoretische Rechtfertigung* des Ranking-Verfahrens wird darin gesehen, daß es den Bedingungen der Willensbildung im Bereich der Politik entspricht, weil das Wesen der Politik gerade darin bestehe, Entscheidungen zwischen wünschenswerten Alternativen treffen zu müssen (Inglehart/Klingemann 1996). In der Politik könne man eben nicht alles auf einmal haben. Oder, wie Inglehart es formuliert: "Politics is above all about choices" (Inglehart 1994: 290). Das Ranking-Verfahren zwingt die Befragten dementsprechend angemessenerweise zur Setzung von Prioritäten (vgl. auch Bauer-Kaase/Kaase 1998: 262). Dies entspreche der Definition gesellschaftlicher Wertorientierungen als internalisierte Konzeptionen des sozial Wünschenswerten, die für das Individuum in Entscheidungssituationen die Funktion von Selektionsstandards erfüllen. Die Messung von Wertorientierungen - so Inglehart (1982: 449) - sei daher auch nur in Konfliktsituationen möglich und sinnvoll, die durch das Ranking-Verfahren adäquat nachgebildet würden. Inglehart spricht vor diesem Hintergrund in seinen Arbeiten auch nicht von *Wertorientierungen*, sondern von *Wertprioritäten*.

2.2 Das Rating-Verfahren

Ein wichtiger *methodischer Vorteil* des Rating-Verfahrens besteht darin, daß es umfragetechnisch leichter zu administrieren und für die Befragten weniger anstrengend ist als das Ranking-Verfahren (Munson/McIntyre 1979: 49). Darüber hinaus ist es ökonomischer in der Durchführung: Empirische Untersuchungen haben gezeigt, daß eine Datenerhebung über Rating-Skalen in der Regel nur ein Drittel der Zeit in Anspruch nimmt, die für ein komplettes Ranking der Wertitems benötigt wird (Munson/McIntyre 1979: 49). Zudem sind Rating-Skalen in Telefoninterviews leichter zu erheben (Alwin/Krosnick 1985: 537).

Ein weiterer Vorteil ist in der Tatsache zu sehen, daß die Datenerhebung über Rating-Skalen einen höheren Informationsgehalt hat als eine Datenerhebung über Ranking (Russell/Gray 1994: 80). Während Rangplätze nur relativ zueinander interpretiert werden können, ermöglicht das Rating-Verfahren auch Aussagen über die absolute Wichtigkeit, die die Befragten den verschiedenen Werten zumessen (Russell/Gray 1980: 80). Rating-Skalen ermöglichen somit auch den interindividuellen Vergleich.

Der wichtigste *methodische Nachteil* des Rating-Verfahrens wird in seiner Anfälligkeit für *Response-Sets*, d.h. inhaltsunabhängiges Antwortverhalten gesehen. In diesem Zusammenhang ist zunächst die sog. *Zustimmungstendenz* der Befragten zu nennen (Hellevik 1994: 293). Diese verursacht ein sog. "end piling", d.h. eine Ballung der vergebenen Skalenwerte am oberen Ende der Bewertungsskala (Munson/McIntyre 1979: 49). Über die verschiedenen Items hinweg läßt sich dann nur eine geringe Varianz der Wichtigkeitszuschreibungen beobachten (Bean/Papadikis 1994: 274; van Deth 1983: 410). Da das Ausmaß der Zustimmungstendenz von Person zu Person variiert, sind die verschiedenen Items in aller Regel untereinander positiv korreliert (Alwin/Krosnick 1985: 537; Krosnick/Alwin 1988: 528). Aufgrund der Tatsache, daß Response-Sets am wahrscheinlichsten bei Niedriggebildeten auftreten, werden außerdem u.U. substantielle positive Zusammenhänge der formalen Bildung mit bestimmten Wertorientierungen verdeckt (Krosnick/Alwin 1988: 530).

Die eben beschriebene Zustimmungstendenz kann im Extremfall so weit gehen, daß die Befragten komplizierte Entscheidungen über die relative Wichtigkeit der vorgegebenen Werte vollständig vermeiden und *alle* Werte als in *gleichem* Maße wünschenswert einstufen. Diesen Effekt bezeichnet man als *Nicht-Differenzierung* (Krosnick/Alwin 1988: 526). Nicht-Differenzierung wird von den Befürwortern des Ranking-Verfahrens als das Ergebnis

unreflektierten Antwortverhaltens bzw. mangelnder Motivation auf seiten der Befragten gedeutet, was die Validität der Messung von Werten über Rating-Skalen in Frage stellt (z.B. Inglehart 1994; Inglehart/Klingemann 1996). Prinzipiell besteht aber auch die Möglichkeit, daß die Befragten wirklich alle abgefragten Werte als gleich wichtig einstufen, die Messung also valide ist (Krosnick/Alwin 1988: 536).

Die *theoretische Rechtfertigung* des Rating-Verfahrens wird darin gesehen, daß es keine trade-offs zwischen Werten erzwingt, wo materiell keine existieren (Alwin/Krosnick 1985: 549). Mit anderen Worten: "Likert-Scaling may be more isomorphic or 'similar in form' to reality than forced rank ordering" (Munson/McIntyre 1979: 49). Die Nicht-Existenz bzw. -Relevanz von Trade-off-Prozessen läßt sich inhaltlich wie folgt begründen: Zum einen implizieren Wertorientierungen als Konzeptionen des sozial Wünschenswerten noch nicht notwendigerweise Prioritätensetzungen (Klein 1995: 211). Vielmehr kommt es erst im politischen Prozeß der Umsetzung dieser Vorstellungen in Politikpräferenzen zu einer solchen Situation. Das Ranking-Verfahren ist also angemessen, wenn es um die Erhebung von Issueprioritäten geht, nicht aber, wenn man die Messung von Konzeptionen des Wünschenswerten intendiert (Bauer-Kaase/Kaase 1998: 262). Zum anderen sind die konkreten Ziele des Inglehart-Index nicht notwendigerweise kompetitiv, es sei denn, man behauptet eine generell eng begrenzte Problemverarbeitungskapazität des politischen Systems. Ansonsten aber läßt sich eine Trade-off-Beziehung zwischen beispielsweise dem Schutz des Rechts auf freie Meinungsäußerung und der Preisniveaustabilität theoretisch kaum begründen (vgl. Bürklin/Klein/Ruß 1996: 520f.).

3. Datenbasis und Operationalisierung

Die Daten, die wir im folgenden auswerten, wurden im Rahmen des Forschungsprojekts "Die Conjoint-Analyse als Instrument der empirischen Wahlforschung" erhoben (vgl. Arzheimer/Klein 1998a). Es handelt sich hierbei um eine zweiwellige schriftliche Panelbefragung, die wir anläßlich der Bürgerschaftswahl 1997 in Hamburg durchgeführt haben. Die Feldphase der ersten Welle startete Ende September 1997, die der zweiten Panelwelle Ende Februar 1998. Die Bruttostichprobe umfaßte 4.000 Personen, die vom Hamburger Amt für zentrale Meldeangelegenheiten sechs Wochen vor der Wahl zufällig aus der wahlberechtigten Bevölkerung ausgewählt worden waren. In der ersten Panelwelle erhielten wir 2.268 ausgefüllte Fragebögen zurück. Unter Berücksichtigung der stichprobenneutralen Ausfälle (Adreßfehler) betrug die Rücklaufquote 58,5 Prozent. Die Haltequote der zweiten Panelwelle betrug 87,1 Prozent, so daß uns insgesamt 1.882 auswertbare Fälle zur Verfügung stehen (zum Ablauf der Datenerhebung vgl. ausführlich Arzheimer/Klein 1998b). Da der für die folgenden Analysen zentrale "hybride" Inglehart-Index im Rahmen eines Split-half-Experiments nur einem zufällig ausgewählten Teil der Zielpersonen vorgelegt wurde, beschränken sich unsere Analysen in diesem Aufsatz auf eine Teilstichprobe von 1.127 Befragten in der ersten und 939 Befragten in der zweiten Panelwelle.

Um die beiden zur Diskussion stehenden Erhebungsformate miteinander vergleichen zu können, haben wir die vier Items des Inglehart-Index zunächst unabhängig voneinander über siebenstufige Ratingskalen bewerten lassen und die Befragten anschließend gebeten, diese Items außerdem in eine Rangordnung zu bringen (vgl. für ein ähnliches Vorgehen bei der 12-Item-Version des Inglehart-Index Bean/Papadikis 1994). Dieses Erhebungsformat machte eine Anpassung des Fragestimulus notwendig. In der Regel wird der Inglehart-Index mit dem Satz "Auch in der Politik kann man nicht alles auf einmal haben" eingeleitet. Da in unserem Fall die Befragten aber zunächst die Items unabhängig voneinander in ihrer Wichtigkeit bewerten sollten, war dieser Einleitungssatz unangemessen. Vor diesem Hintergrund haben wir uns für eine modifizierte Version des in den Eurobarometern der EU-Kommission verwendeten Fragestimulus entschieden (vgl. Krebs/Hofrichter 1989: 62). "Es wird in Deutschland gegenwärtig viel darüber geredet, welche Ziele die Politik in den nächsten Jahren verfolgen sollte. Hier sind einige Ziele genannt, die man in der Politik anstreben kann. Bitte geben Sie für jedes Ziel an, für wie wichtig Sie es halten, daß dieses Ziel in der Politik verfolgt wird." Im

Anschluß daran haben wir dann weiter gefragt: "Einmal angenommen, Sie müßten sich zwischen diesen vier Zielen entscheiden. Welches Ziel wäre Ihnen dann am wichtigsten, welches am zweitwichtigsten und welches am dritt wichtigsten?" Es ist nicht zu erwarten, daß dieser veränderte Fragestimulus nennenswerte Verzerrungen der Messungen induziert, da die Befunde von Krebs und Hofrichter darauf hindeuten, daß unterschiedliche Frageformulierungen im Rahmen des Inglehart-Index keinen Einfluß auf die Meßergebnisse besitzen (Krebs/Hofrichter 1989).

4. Der empirische Befund

4.1 Die Wichtigkeit der Ziele des Inglehart-Index

Betrachten wir zunächst die absolute bzw. relative Wichtigkeit, die die Befragten den vier Zielen des Inglehart-Index zuweisen. Bei der absoluten Wichtigkeitszuschreibung über Rating-Skalen läßt sich deutlich die oben beschriebene Zustimmungstendenz beobachten (vgl. Tab. 1). Bei allen vier Zielen ist der am häufigsten vergebene Skalenwert (Modus) die "7". Der Prozentanteil dieses Skalenwertes liegt bei den verschiedenen Zielvorgaben zwischen 34,6 und 60,6 Prozent, die zugehörigen arithmetischen Mittelwerte liegen im Bereich zwischen 5,5 und 6,2. Die Befragten halten also alle vier Ziele für wichtig, es sind aber durchaus Abstufungen in den absoluten Wichtigkeiten zu erkennen. Als wichtigstes politisches Ziel wird im Aggregat der Schutz des Rechts auf freie Meinungsäußerung angesehen, gefolgt von der Aufrechterhaltung von Ruhe und Ordnung und einem größeren Bürgereinfluß auf Regierungsentscheidungen. Der Kampf gegen steigende Preise wird mit deutlichem Abstand als am wenigsten wichtig eingestuft.

Tab. 1: Die absolute Wichtigkeit der Ziele des Inglehart-Index



	<i>Prozentuale Häufigkeit der Skalenwerte</i>							Modus	
	1	2	3	4	5	6	7		
A Ruhe und Ordnung	2,0	3,5	4,9	7,6	11,0	18,3	52,6	5,9	7
B Mehr Bürgereinfluß	0,9	1,1	3,6	9,7	17,5	23,5	43,8	5,9	7
C Kampf gegen steigende Preise	1,7	2,6	5,3	12,2	22,3	21,2	34,6	5,5	7
D Freie Meinungsäußerung	0,8	1,5	2,5	7,0	8,6	19,1	60,5	6,2	7

Beim Ranking-Verfahren wird auf individueller Ebene die Zuweisung relativer Wichtigkeiten durch das Abfrageformat erzwungen. Als wichtigstes politisches Ziel wird dabei am häufigsten die Aufrechterhaltung von Ruhe und Ordnung genannt; 47,5 Prozent der Befragten setzen dieses Ziel auf den ersten Rang (vgl. Tab. 2). Es folgen das Recht auf freie Meinungsäußerung, das von 27,7 Prozent der Befragten als am wichtigsten angesehen wird, der Bürgereinfluß auf Regierungsentscheidungen (17,5 Prozent) sowie der Kampf gegen steigende Preise (7,3 Prozent). Je nachdem, ob man die über Rating oder die über Ranking erhobenen Daten heranzieht, unterscheidet sich also im Aggregat die relative Wichtigkeit der verschiedenen Zielvorstellungen. Im ersten Fall erweist sich das Recht auf freie Meinungsäußerung als am wichtigsten, im zweiten die Aufrechterhaltung von Ruhe und Ordnung. Dies widerspricht dem in der Literatur wiederholt berichteten Befund, daß Ratings und Rankings im Aggregat ähnliche Präferenzordnungen produzieren (vgl. u.a. Feather 1973; Moore 1975; Alwin/Krosnick 1985).

Tab. 2: Die relative Wichtigkeit der Ziele des Inglehart-Index (in Prozent)

	<i>Rangplatz</i>			mittlerer Rangplatz
	1	2	3	
A Ruhe und Ordnung	47,5	20,4	14,9	2,0
B Mehr Bürgereinfluß	17,5	33,6	29,1	2,5
C Kampf gegen steigende Preise	7,3	24,6	29,9	3,0
D Freie Meinungsäußerung	27,7	21,4	26,0	2,5
Modus	A	B	C	

Eine mögliche Erklärung für diesen Unterschied könnte darin bestehen, daß auch beim Ranking-Verfahren Response-Sets auftreten. So läßt sich in Tab. 2 erkennen, daß als wichtigstes Ziel am häufigsten "A" genannt wird, als zweitwichtigstes Ziel am häufigsten "B" und als drittwichtigstes Ziel am häufigsten "C". Könnte es also bei der Rangordnungsaufgabe nicht eine Strategie der Befragten sein, unreflektiert einfach die Rangfolge "ABC" zu nennen, ohne sich Gedanken über die Inhalte der Items zu machen? Im nächsten Abschnitt wollen wir dieses Problem genauer untersuchen.

4.2 Die Existenz von Response-Sets

Beginnen wir mit der Verteilung der Antworten auf den Rating-Skalen. In Tab. 3 ist die Häufigkeit individuell gleicher Skalenpunktwahl dargestellt. Insgesamt zeigt sich, daß höhere Skalenpunkte mit einer sehr viel größeren Wahrscheinlichkeit gewählt werden als niedrigere Skalenpunkte. Als problematisch muß betrachtet werden, daß 19,1 Prozent der Befragten 4 mal den Skalenpunkt 7 wählen, also in ihrer absoluten Wichtigkeitszuschreibung überhaupt nicht zwischen den vier Items differenzieren. Weitere 1,6 Prozent der Befragten vergeben 4 mal einen der anderen Skalenwerte, so daß im Ergebnis 20,7 Prozent der Befragten eine Variationsbreite der Skalenausnutzung von 0 aufweisen. Diese Gruppe von Befragten wollen wir im folgenden *Nicht-Differenzierer* nennen.

Tab. 3: Häufigkeit individuell gleicher Skalenpunktwahl bei den Rating-Skalen

	<i>Prozentuale Häufigkeit der Wahl</i>				
	0 mal	1 mal	2 mal	3 mal	4 mal
Skalenpunkt 1	95,9	3,3	0,3	0,2	0,2
Skalenpunkt 2	92,1	7,1	0,6	0,1	0,0
Skalenpunkt 3	86,2	11,6	2,0	0,1	0,1
Skalenpunkt 4	69,4	25,1	5,1	0,3	0,1
Skalenpunkt 5	56,3	30,6	10,6	2,3	0,2
Skalenpunkt 6	47,6	29,5	17,0	4,8	1,0
Skalenpunkt 7	19,0	24,2	22,4	15,4	19,1

Wenn es sich bei den Nicht-Differenzierern tatsächlich um Befragte handelt, die den kognitiven Aufwand bei der Beantwortung des Fragebogens durch eine Ja-Sage-Strategie zu reduzieren versuchen, dann sollte der Anteil der Nicht-Differenzierer unter Niedriggebildeten höher sein als unter Hochgebildeten, da letztere eher das kognitive Rüstzeug besitzen, die Rating-Aufgabe differenziert zu erfüllen (Krosnick/Alwin 1988: 529f). Tatsächlich zeigt sich, daß unter den Hauptschulabsolventen 37,5 Prozent den Nicht-Differenzierern zuzurechnen sind, während dieser Anteil bei den Realschulabsolventen nur 21,4 Prozent und bei den Abiturienten gar nur 6,9 Prozent beträgt. Außerdem steht zu erwarten, daß Response-Sets um so eher auftreten, je geringer das Interesse am Thema der Befragung ist. Und in der Tat: Während unter den Personen mit einem hohen oder sehr hohen Politikinteresse nur 17,8 Prozent der Befragten nicht zwischen den Items differenzieren, ist der Anteil der Nicht-Differenzierer bei den Befragten mit einem geringen oder sehr geringen politischen Interesse mit 32,4 Prozent fast doppelt so hoch.

Die Befürworter des Ranking-Verfahrens argumentieren nun, daß diejenigen Befragten, die im Rahmen des Rating-Formats schwierigen Entscheidungen durch Nicht-Differenzierung ausweichen, in der durch das Ranking-Format definierten Forced-choice-Situation gezwungen werden, reflektierte Urteile abzugeben. Diese Erwartung ist nicht unbedingt plausibel. Wenn Befragte mit einem nur gering ausgeprägten kognitiven Differenzierungsvermögen bereits durch das Rating-Verfahren überfordert werden, so kann vermutet werden, daß sie durch das komplexere Ranking erst recht überfordert sind. Als erster Beleg für eine solche Interpretation kann angeführt werden, daß unter den Nicht-Differenzierern 7,6 Prozent die Rangordnungsaufgabe vollständig verweigern, während dies von den restlichen Befragten nur 2,7 Prozent tun. Dieser Unterschied ist statistisch signifikant ($p < 0.01$).

Darüber hinaus sollte man erwarten können, daß Nicht-Differenzierer mit einer höheren Wahrscheinlichkeit auch bei der Erfüllung der Ranking-Aufgabe stereotyp antworten, also beispielsweise das erstgenannte Ziel auch auf den ersten Platz setzen. Wir haben deshalb im Rahmen unserer Panelbefragung ein Methodenexperiment durchgeführt, das es uns ermöglicht, solche Response-Sets zu identifizieren. Dazu haben wir in der zweiten Panelwelle im Rahmen eines Split-Half-Experiments der einen Hälfte der Befragten wieder den ursprünglichen hybriden Inglehart-Index vorgelegt, während wir der anderen Hälfte der Befragten eine Version mit rotierten Itemvorgaben vorlegten, bei dem jeweils das erste und das zweite sowie das dritte und das vierte Ziel miteinander vertauscht wurden. Wenn sich hierbei nachweisen läßt, daß die Reihenfolge der Nennung der Ziele einen Einfluß auf deren Chance hat, als wichtigstes politisches Ziel genannt zu werden, dann ist dies als Beleg für die Existenz von Response-Sets auch im Rahmen des Ranking-Verfahrens zu werten. Unter den Nicht-Differenzierern der ersten Panelwelle sollte dieser Plazierungseffekt besonders deutlich ausfallen.

Die Ergebnisse dieses Methodenexperiments sind in Tab. 4 dokumentiert. Als wichtigster Befund ist festzuhalten, daß ein signifikanter Plazierungseffekt auftritt: Das Ziel "Mehr Einfluß der Bürger auf die Entscheidungen der Regierung" wird von 13,9 Prozent der Befragten als wichtigstes politisches Ziel genannt, wenn es im Fragebogen an zweiter Stelle steht, hingegen von 21,3 Prozent der Befragten, wenn es die Liste der politischen Ziele anführt. Betrachtet man das Antwortverhalten der Differenzierer und Nicht-Differenzierer der 1. Panelwelle getrennt voneinander, so zeigt sich, daß der Plazierungseffekt bei den Nicht-Differenzierern stärker ausfällt. Bei ihnen erhöht sich der Anteil der Befragten, der das Ziel "Mehr Bürgereinfluß" auf den ersten Rangplatz setzt, nach der Itemrotation von 12,1 auf 25 Prozent, während die entsprechenden Prozentsätze beim Rest der Befragten 14,6 und 19,8 betragen. Die eben beschriebenen Plazierungseffekte sind in den durch die ersten beiden Zeilen von Tab. 4 definierten Teiltabellen jeweils auf dem 5%-Niveau statistisch signifikant.

Tab. 4: Response-Sets im Rahmen des Ranking-Verfahrens

1. Rangplatz	alle Befragten der 2. Panelwelle		Differenzierer der 1. Panelwelle		Nicht-Differenzierer der 1. Panelwelle	
	original	rotiert	original	rotiert	original	rotiert
Ruhe und Ordnung	51,3	44,8	47,1	41,7	68,1	57,6
Mehr Bürgereinfluß	13,9	21,3	14,6	19,8	12,1	25,0
Kampf gegen steigende Preise	8,3	8,0	8,1	7,8	8,8	8,7
Freie Meinungsäußerung	26,5	25,9	30,3	30,7	11,0	8,7
N=	460	451	357	348	91	92
Zusammenhang (Cramer's V)						
(für die beiden erstgenannten Ziele)	V= .12 (p= .003)		V= .10 (p= .046)		V= .18 (p= .027)	

Anmerkung: Eintragungen sind der Prozentanteil der Befragten, die das jeweilige politische Ziel auf den ersten Rangplatz setzen.

Die Nicht-Differenzierer müssen unter den Bedingungen des Ranking-Verfahrens aber nicht notwendigerweise eine systematische Response Strategie wählen. Mindestens genauso plausibel erscheint die Annahme, daß ein Teil der Befragten die Rangplätze zufällig vergibt, ohne sich Gedanken über den konkreten Inhalt der verschiedenen Ziele zu machen. Wenn dieses Muster vorliegen würde, dann müßte die Test-Retest-Stabilität der über das Ranking-Verfahren gemessenen Wertorientierungen in der Gruppe der Nicht-Differenzierer sehr viel niedriger liegen als in der Gruppe der Differenzierer. Wir haben daher auf der Grundlage der vergebenen Rangplätze die gängige Einteilung in Materialisten, Mischtypen und Postmaterialisten vorgenommen und die Stabilität dieser Zuordnung über die beiden Panelwellen hinweg analysiert. Um die eben diskutierten Plazierungseffekte auszuschalten, haben wir uns dabei auf die Gruppe der Befragten beschränkt, die in beiden Panelwellen den unrotierten Inglehart-Index vorgelegt bekam. Die Stabilität der Zuordnung haben wir über das ordinale Assoziationsmaß γ

gemessen. Für die Gruppe der Differenzierer betrug dieses $\gamma = 0,84$, bei den Nicht-Differenzierern hingegen nur $\gamma = 0,57$.

4.3 Die Zuordnung der Befragten zu den Wertetypen

Wenden wir uns abschließend der Frage nach der Verteilung der unterschiedlichen Wertetypen in Abhängigkeit vom zugrundeliegenden Erhebungsverfahren zu. Bei den Ranking-Daten erfolgt die Zuordnung der Befragten zu den drei Wertetypen nach den altbekannten Regeln. Bei den Rating-Daten haben wir all diejenigen Befragten als Postmaterialisten klassifiziert, die *beiden* postmaterialistischen Zielen jeweils höhere Wichtigkeiten zuschreiben als den beiden materialistischen Zielen. Bei den Materialisten sind wir analog verfahren. Als Mischtypen haben wir alle übrigen Befragten klassifiziert. Diese Art der Skalenbildung lehnt sich insofern eng an die Konzeption Ingleharts an, als sie die vergebenen Skalenwerte nur relativ zueinander interpretiert.

Tab. 5 zeigt, wie sich die empirischen Randverteilungen der zwei eben beschriebenen alternativen "Inglehart-Indizes" darstellen. Im Rahmen des originalen Inglehart-Index werden 19,4 Prozent der Befragten als Materialisten, 60,6 Prozent als Mischtypen und 20,1 Prozent als Postmaterialisten ausgewiesen. Die entsprechenden Anteile auf der Grundlage der Rating-Skalenwerte betragen 5,2 Prozent, 82,1 Prozent und 12,7 Prozent. Daß der Anteil der Mischtypen im Rahmen des Rating-Formates deutlich höher liegt, entspricht den Erwartungen: Da das Rating-Format keine "Forced-choice"-Situation darstellt, ist es naheliegend, daß mehr Befragte gemischte Wertpräferenzen zu erkennen geben als im Rahmen des Ranking-Verfahrens.

Tab. 5: Die Verteilung der Wertetypen in Abhängigkeit vom Erhebungsformat (in Prozent)

	Ranking	Rating
Materialist	19,4	5,2
Mischtyp	60,6	82,1
Postmaterialist	20,1	12,7
N =	1037	1080

Aufschlußreich ist darüber hinaus der Zusammenhang zwischen den beiden Varianten des Inglehart-Index. A priori kann man die Erwartung formulieren, daß bei konsistentem Antwortverhalten der Befragten kein Wechsel zwischen den Extremkategorien zu beobachten sein sollte, d.h. niemand, der im Rahmen des Ranking-Verfahrens Postmaterialist ist, sollte im Rahmen des Rating-Verfahrens plötzlich zum Materialisten mutieren und umgekehrt. Wie Tab. 6 zeigt, entspricht diese Erwartung der Realität. Darüber hinaus sollte von den im Rahmen des Ranking-Verfahrens als Mischtyp klassifizierten Befragten der allergrößte Teil auch im Rahmen des Ratings wieder in die Mischkategorie fallen, da Befragte, die bereits unter Forced-choice-Bedingungen differenzierte Prioritäten erkennen lassen, dies im Rahmen des unrestringierten Rating-Formates erst recht tun sollten. Und tatsächlich werden nur 4,8 Prozent der Mischtypen des originalen Inglehart-Index im Rahmen des Rating-Formates einem der beiden reinen Wertetypen zugeordnet.

Tab. 6: Die Beziehung zwischen dem über das Ranking-Verfahren erhobenen und dem über das Rating-Verfahren erhobenen Inglehart-Index

	Ranking-Verfahren			
Rating-Verfahren	Materialist	Mischtyp	Postmaterialist	Σ
Materialist	21,6 4,2	1,6 1,0	0,0 0,0	5,2
Mischtyp	78,4 15,1	95,2 57,8	44,7 8,9	81,8
Postmaterialist	0,0 0,0	3,2 1,9	55,3 11,1	13,0

Σ	100,0	19,3	100,0	60,7	100,0	20,0	100,0
---	-------	------	-------	------	-------	------	-------

Bedeutende Abweichungen zwischen den beiden unterschiedlichen Erhebungsverfahren treten nur dort auf, wo man sie auch erwartet: Von den reinen Wertetypen des originalen Inglehart-Index fallen im Rahmen des Rating-Formates 55,3 Prozent (Postmaterialisten) bzw. 78,4 Prozent (Materialisten) in die Kategorie mit gemischten Werteorientierungen. Nur an dieser Stelle machen Unterschiede zwischen den beiden Erhebungsverfahren auch theoretisch Sinn. Wer beispielsweise im Rahmen des Ranking-Verfahrens zwei materialistische Ziele an die Spitze seiner Prioritätenliste gesetzt hat, muß den postmaterialistischen Zielen notgedrungenerweise eine geringere relative Wichtigkeit zuweisen, auch wenn er sie eigentlich als gleich wichtig ansieht. Im Rahmen eines weniger restriktiven Erhebungsverfahrens besteht diese Notwendigkeit hingegen nicht.

Insgesamt handelt es sich bei den eben beschriebenen "Wechslern" um 24,1 Prozent aller Befragten. Dieser Anteil entspricht in etwa dem Anteil der weiter oben beschriebenen Nicht-Differenzierer. Dies legt die Vermutung nahe, daß diese beiden Gruppen weitgehend identisch sind, daß also die Personen, die im Rahmen des Rating-Verfahrens nicht zwischen den vier Zielen differenzieren und deshalb dort dem Mischtyp zugeordnet werden, aufgrund des Forced-choice-Formates des Rankings nun zur Offenbarung differenzierter Präferenzen gezwungen werden. Empirisch zeigt sich aber, daß von den "Wechslern" nur 29,9 Prozent zur Gruppe der "Nicht-Differenzierer" gehören. Zwar liegt dieser Anteil signifikant höher als bei den restlichen Befragten (16,9 Prozent), doch sind 70,1 Prozent der "Wechsler" offensichtlich in der Lage, auf den Rating-Skalen differenziert zu antworten. Dies läßt sich dahingehend interpretieren, daß die abweichende Randverteilung des auf der Grundlage der Rating-Daten definierten Inglehart-Index in der Hauptsache darauf zurückzuführen ist, daß das Ranking den Befragten Restriktionen auferlegt, die ihnen die Offenbarung ihrer "wahren" Präferenzen unmöglich machen. Als weiterer Beleg für eine solche Interpretation läßt sich außerdem anführen, daß auch unter den Abiturienten 21,6 Prozent der Befragten zu den Wechslern gehören. Diese Befragten sind kognitiv in der Regel durchaus differenzierungsfähig und antworten auch auf den Rating-Skalen nicht stereotyp. Offensichtlich kann diese Personengruppe den vollen Umfang ihrer Differenzierungsfähigkeit im Rahmen des restringierten Ranking nicht adäquat zum Ausdruck bringen.

4.4 Die Zusammenhänge mit dem Alter und der Bildung

Die Entscheidung für eines der beiden untersuchten Erhebungsverfahren bewirkt deutliche Unterschiede in den substantiellen Befunden, die man erzielt. Dies zeigt sich, wenn man die beiden Varianten des Inglehart-Index zu den wichtigsten Determinanten gesellschaftlicher Wertorientierungen, dem Alter und der Bildung, in Beziehung setzt (vgl. Tab. 12). Beim Ranking-Verfahren stehen die reinen Wertetypen jeweils in einem deutlichen Zusammenhang mit diesen Variablen: der Anteil der Materialisten steigt mit höherem Alter und geringerer Bildung an, während gleichzeitig der Anteil der Postmaterialisten sinkt. Demzufolge zeigen sich auch starke Zusammenhänge des Alters und der Bildung mit der von Inglehart in seinen Analysen als abhängige Variable verwendeten Prozentpunktdifferenz zwischen dem Anteil der Postmaterialisten und dem Anteil der Materialisten (PPM).

Tab. 12: Die Zusammenhänge der beiden Varianten des Inglehart-Index mit der Bildung und dem Alter

	<i>formale Bildung</i>			<i>Alter in Jahren</i>				
	Haupt	Real	Abi	18-24	25-34	35-49	50-64	65+
Inglehart-Index (Ranking)								
Materialist	26,2	21,6	13,3	8,1	16,8	15,6	21,7	31,5
Mischtyp	63,3	63,5	56,4	59,3	58,8	54,4	65,1	63,5

Postmaterialist	10,5	14,9	30,3	32,6	24,4	30,0	13,2	5,1
PPM	-15,7	-6,7	17,0	24,5	7,6	14,4	-8,5	-26,4
Inglehart-Index (Rating)								
Materialist	4,4	6,3	5,2	4,7	4,5	7,1	5,1	4,6
Mischtyp	90,5	82,9	75,0	77,6	74,6	75,4	87,5	92,8
Postmaterialist	5,1	10,9	19,8	17,6	20,9	17,5	7,4	2,6
PPM	0,7	4,6	14,6	12,9	16,4	10,4	2,3	-2,0

Anmerkungen: PPM = Prozentpunktdifferenz zwischen dem Anteil der Postmaterialisten und dem Anteil der Materialisten.

Anders stellt sich die Situation dar, wenn man den auf der Grundlage der Rating-Skalen konstruierten Inglehart-Index betrachtet. Hier besteht weder ein Zusammenhang zwischen dem Alter und dem Anteil der Materialisten noch zwischen der formalen Bildung und dem Anteil der Materialisten. Diese Zusammenhänge kommen bei der Analyse der Ranking-Daten offensichtlich nur dadurch zustande, daß Materialismus und Postmaterialismus nicht unabhängig voneinander gemessen werden. Da Alter und Bildung im Falle der Rating-Daten jeweils nur mit dem Anteil der Postmaterialisten korrelieren, schwächen sich auch ihre Zusammenhänge mit dem PPM erkennbar ab (vgl. Tab. 12). Der Wertewandel erscheint so deutlich weniger gravierend als bei den über das Ranking-Verfahren erhobenen Daten. Auch die von Inglehart vorgeschlagene theoretische Erklärung des Wertewandels wird bei der Analyse der Rating-Daten fragwürdig, denn der Trend des Wertewandels geht in diesem Fall nicht von den reinen Materialisten hin zu den reinen Postmaterialisten, sondern von den Mischtypen hin zu den reinen Postmaterialisten.

5. Zusammenfassung und Schlußfolgerungen

Der dem Ranking-Verfahren gemeinhin zugeschriebene methodische Vorteil, die Befragten zu reflektierten Urteilen zu zwingen, läßt sich empirisch nicht bestätigen. Es existiert vielmehr ein signifikanter Einfluß der Reihenfolge, in der die Wertitems präsentiert werden, auf die von den Befragten offenbarte Präferenzordnung. Dieser Einfluß ist bei Personen, die im Rahmen des Rating-Verfahrens allen Werten die gleiche Wichtigkeit zuschreiben (Nicht-Differenzierer), besonders stark. In dieser Personengruppe liegt zudem die Test-Retest-Stabilität des über ein Rangordnungsverfahren erhobenen Inglehart-Index deutlich niedriger als bei den übrigen Befragten.

Diese Befunde lassen sich auf zweierlei Art und Weise deuten: Entweder antworten die Nicht-Differenzierer sowohl auf die Rating- als auch auf die Ranking-Fragen unreflektiert. Der beim Rating auftretende Nicht-Differenzierungs-Effekt könnte dann nicht länger als Argument für die Überlegenheit des Ranking angeführt werden. Oder aber die Nicht-Differenzierer messen wirklich allen Zielen die gleiche Bedeutung zu. Dann könnten diese Personen ihre wahren Präferenzen im Rahmen des Ranking-Formates nicht adäquat zum Ausdruck bringen und würden deshalb dort stereotyp oder zufällig antworten (Krosnick/Alwin 1988: 536). Im ersten Fall hätte das Ranking-Verfahren also keine Vorteile gegenüber dem Rating, im zweiten Fall wäre es ihm sogar unterlegen. Daß der zweite Fall der wahrscheinlichere ist, zeigt sich daran, daß auch 20 Prozent der Hochgebildeten gemischte Wertorientierungen offenbaren, wenn das unrestringierte Rating-Format verwendet wird, während sie im Rahmen des restriktiveren Ranking-Verfahrens "reine" Wertorientierungen zu erkennen geben. Unserer Ansicht nach spricht somit alles dafür, sich bei der Erhebung von Wertorientierungen für das Rating-Verfahren zu entscheiden, da dieses auf die Einschränkung individueller Freiheitsgrade der Bewertung verzichtet (Bürklin/Klein/Ruß 1996: 521). Darüber hinaus ermöglicht es außerdem die Anwendung komplexer statistischer Analyseverfahren sowie eine unverzerrte Bestimmung der Dimensionalität des Werteraums.

Zur Begründung der Überlegenheit des Ranking-Verfahrens lassen sich also bestenfalls noch theoretische Argumente ins Feld führen. Daß diese in der Regel nicht besonders überzeugend sind, haben wir weiter oben bereits ausgeführt. Doch selbst wenn man diese theoretischen Argumente teilt, ergibt sich daraus noch nicht notwendigerweise, daß Rating-Skalen unangemessen sind. "It is not difficult to justify assessing priorities, but that certainly does not automatically justify the use of ipsative measures. As indicated, every set of scored items reveals an ordering of those items, provided some variance is observed in the scores (to break ties). The only reason why ipsative measures are selected so frequently when priorities are relevant seems to be the expectation that insufficient variance will be obtained when independent evaluations are asked. However, that is an empirical and not a theoretical consideration" (van Deth 1983: 411).

Zieht man schließlich in Betracht, daß nicht nur das von Inglehart verwendete Erhebungsformat, sondern auch sein Fragestimulus und seine Items defizitär sind (vgl. zuletzt Bauer-Kaase/Kaase 1998; Clarke et al. 1998), dann kann man mit Bauer-Kaase und Kaase abschließend nur konstatieren, "daß die gängige Operationalisierung der Materialismus-Postmaterialismus-Dimension in einer Weise defekt ist, die eine Neuorientierung unabweisbar macht" (1998: 272).

Literaturverzeichnis

- Alwin, Duane F. / Krosnick, Jon A., 1985: The Measurement of Values in Surveys: A Comparison of Ratings and Rankings. In: Public Opinion Quarterly, 49, S. 535-552.*
- Arzheimer, Kai / Klein, Markus, 1998a: Die Conjoint-Analyse als Instrument der empirischen Wahlforschung. Papier für die Tagung "Wahlen und Wahlforschung: Kontinuität und Wandel" des Arbeitskreises "Wahlen und politische Einstellungen" der DVPW vom 14. bis 15. Mai 1998 in Mainz. [www.uni-mainz.de/~arzheim/conjoint/Conjoint.html]*
- Arzheimer, Kai / Klein, Markus, 1998b: Die Wirkung materieller Incentives auf den Rücklauf einer schriftlichen Panelbefragung. In: ZA-Information, 43, S. 6-31.*
- Bauer-Kaase, Petra / Kaase, Max, 1998: Werte und Wertewandel - ein altes Thema und eine neue Facette. In: Galler, Heinz P. / Wagner, Gerd (Hrsg.): Empirische Forschung und wirtschaftspolitische Beratung. Festschrift für Hans-Jürgen Krupp zum 65. Geburtstag. Frankfurt/Main; New York: Campus Verlag (=Reihe "Wirtschaftswissenschaft"; Bd. 38); S. 256-274.*
- Bean, Clive / Papadikis, Elim, 1994a: Polarized Priorities or Flexible Alternatives? Dimensionality in Inglehart's Materialism-Postmaterialism-Scale. In: International Journal of Public Opinion Research, 6, S. 264-288.*
- Bean, Clive / Papadikis, Elim, 1994b: Polarized Priorities and Flexible Alternatives: Response to Inglehart and Hellevik. In: International Journal of Public Opinion Research, 6, S. 295-297.*
- Bürklin, Wilhelm / Klein, Markus / Ruß, Achim, 1994: Dimensionen des Wertewandels. Eine empirische Längsschnittanalyse zur Dimensionalität und der Wandlungsdynamik gesellschaftlicher Wertorientierungen. In: Politische Vierteljahresschrift, 35, S. 579-606.*
- Bürklin, Wilhelm / Klein, Markus / Ruß, Achim, 1996: Postmaterieller oder anthropozentrischer Wertewandel? Eine Erwiderung auf Ronald Inglehart und Hans-Dieter Klingemann. In: Politische Vierteljahresschrift, 37, S. 517-536.*
- Clarke, Harold D. / Kornberg, Allan / McIntyre, Chris with Bauer-Kaase, Petra / Kaase, Max, 1998: Measuring Value Change and the "Conversation in Context" Critique: A Note with new Experimental Evidence. Unpublished Paper.*
- Feather, N. T., 1973: The Measurement of Values: Effects of Different Assessment Procedures. In: Australian Journal of Psychology, 25, S. 221-231.*
- Flanagan, Scott C., 1987: Value Change in Industrial Societies. In: American Political Science Review, 81, S. 1303-1319.*
- Hellevik, Ottar, 1994: Measuring Cultural Orientations: Rating versus Ranking. In: International Journal of Public Opinion Research, 6, S. 289-297.*

- Hicks, L. E.*, 1970: Some properties of ipsative, normative and forced-choice normative measurement. In: *Psychological Bulletin*, 74, S. 167-184.
- Inglehart, Ronald*, 1982: Changing Values in Japan and the West. In: *Comparative Political Studies*, 14, S. 445-480.
- Inglehart, Ronald*, 1994: "Polarized Priorities or Flexible Alternatives? Dimensionality in Inglehart's Materialism-Postmaterialism-Scale": A Comment. In: *International Journal of Public Opinion Research*, 6, S. 289-297.
- Inglehart, Ronald / Klingemann, Hans-Dieter*, 1996: Dimensionen des Wertewandels. Theoretische und methodische Reflexionen anlässlich einer neuerlichen Kritik. In: *Politische Vierteljahresschrift*, 37, S. 319-340.
- Jackson, David J. / Alwin, Duane F.*, 1980: The Factor Analysis of Ipsative Measures. In: *Sociological Methods & Research*, 9, S. 218-238.
- Klein, Markus*, 1995: Wieviel Platz bleibt im Prokrustesbett? Wertewandel in der Bundesrepublik Deutschland zwischen 1973 und 1992 gemessen anhand des Inglehart-Index. In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 47, S. 207-230.
- Krebs, Dagmar / Hofrichter, Jürgen*, 1989: Materialismus-Postmaterialismus: Effekte unterschiedlicher Frageformulierungen bei der Messung des Konzepts von Inglehart. In: *ZUMA-Nachrichten*, 24, S. 60-72.
- Krosnick, Jon A. / Alwin, Duane F.*, 1988: A Test of the Form-Resistant Correlation Hypothesis. Ratings, Rankings, and the Measurement of Values. In: *Public Opinion Quarterly*, 52, S. 526-538.
- Maag, Gisela*, 1991: *Gesellschaftliche Werte. Strukturen, Stabilität und Funktion*. Opladen: Westdeutscher Verlag (=Beiträge zur sozialwissenschaftlichen Forschung Bd. 120).
- Moore, Michael*, 1975: Rating versus Ranking in the Rokeach Value Survey: An Israeli comparison. In: *European Journal of Social Psychology*, 5, S. 405-408.
- Munson, J. Michael / McIntyre, Shelby H.*, 1979: Developing Practical Procedures for the Measurement of Personal Values in Cross-Cultural Marketing. In: *Journal of Marketing Research*, 16, S. 48-52.
- Rankin, William L. / Grube, Joel W.*, 1980: A comparison of ranking and rating procedures for value system measurement. In: *European Journal of Social Psychology*, 10, S. 233-246.
- Reynolds, Thomas J. / Jolly, James P.*, 1980: Measuring Personal Values: An Evaluation of Alternative Methods. In: *Journal of Marketing Research*, 17, S. 531-536.
- Rockeach, Milton* 1973: *The Nature of Human Values*. New York: The Free Press.
- Russell, P.A. / Gray, C.D.*, 1994: Ranking or Rating? Some data and their implications for the measurement of evaluative response. In: *British Journal of Psychology*, 85, S. 79-92.
- Strobel, Martin*, 1979: Evaluation einiger moderner statistischer Analyseverfahren zur Messung von Werthaltungen. In: Klages, Helmut / Kmiecik, Peter (Hrsg.): *Wertewandel und gesellschaftlicher Wandel*. Frankfurt/New York: Campus, S. 485-504.
- van Deth, Jan W.*, 1983: Ranking the Ratings: The Case of Materialist and Postmaterialist Value Orientations. In: *Political Methodology*, 11, S. 63-79.
- van Deth, Jan W. / Scarbrough, Elinor*, 1995: The Concept of Values. In: Dies. (Hrsg.): *The Impact of Values. Beliefs in Government Volume Four*. Oxford: Oxford University Press, S. 21-47.

Zusammenfassung:

Über die Frage, ob Ranking-Prozeduren wirklich besser für die Messung von Wertorientierungen geeignet sind als Rating-Skalen, gibt es in der Empirischen Sozialforschung eine andauernde Diskussion. Um diese Debatte mit mehr empirischer Evidenz anzureichern, haben wir ein Methodenexperiment in eine zweiwellige schriftliche Panelbefragung integriert: In der ersten Panelwelle wurden die Befragten gebeten, die vier Items des Inglehart-Index sowohl unabhängig voneinander in ihrer Wichtigkeit zu bewerten als auch rangzuordnen. In der zweiten Panelwelle haben wir diese Bitte wiederholt, dabei allerdings der Hälfte der Respondenten einen Fragebogen

vorgelegt, in dem die Reihenfolge der Items vertauscht war. Mit Hilfe dieses experimentellen Designs können wir zeigen, daß ein Zusammenhang zwischen der Reihenfolge der Items im Fragebogen und den von den Befragten vergebenen Rangpositionen besteht. Dies gilt in besonderem Maße für Befragte, die in der ersten Panelwelle im Rahmen des Rating-Verfahrens allen Items dieselbe Wichtigkeit zugeschrieben hatten. Das Ranking-Verfahren ist also für Resonse-Sets ebenso anfällig wie Rating-Skalen. Vor dem Hintergrund dieser Befunde besteht kein Grund für die Annahme, daß das Ranking dem Rating in methodischer Hinsicht überlegen sei.